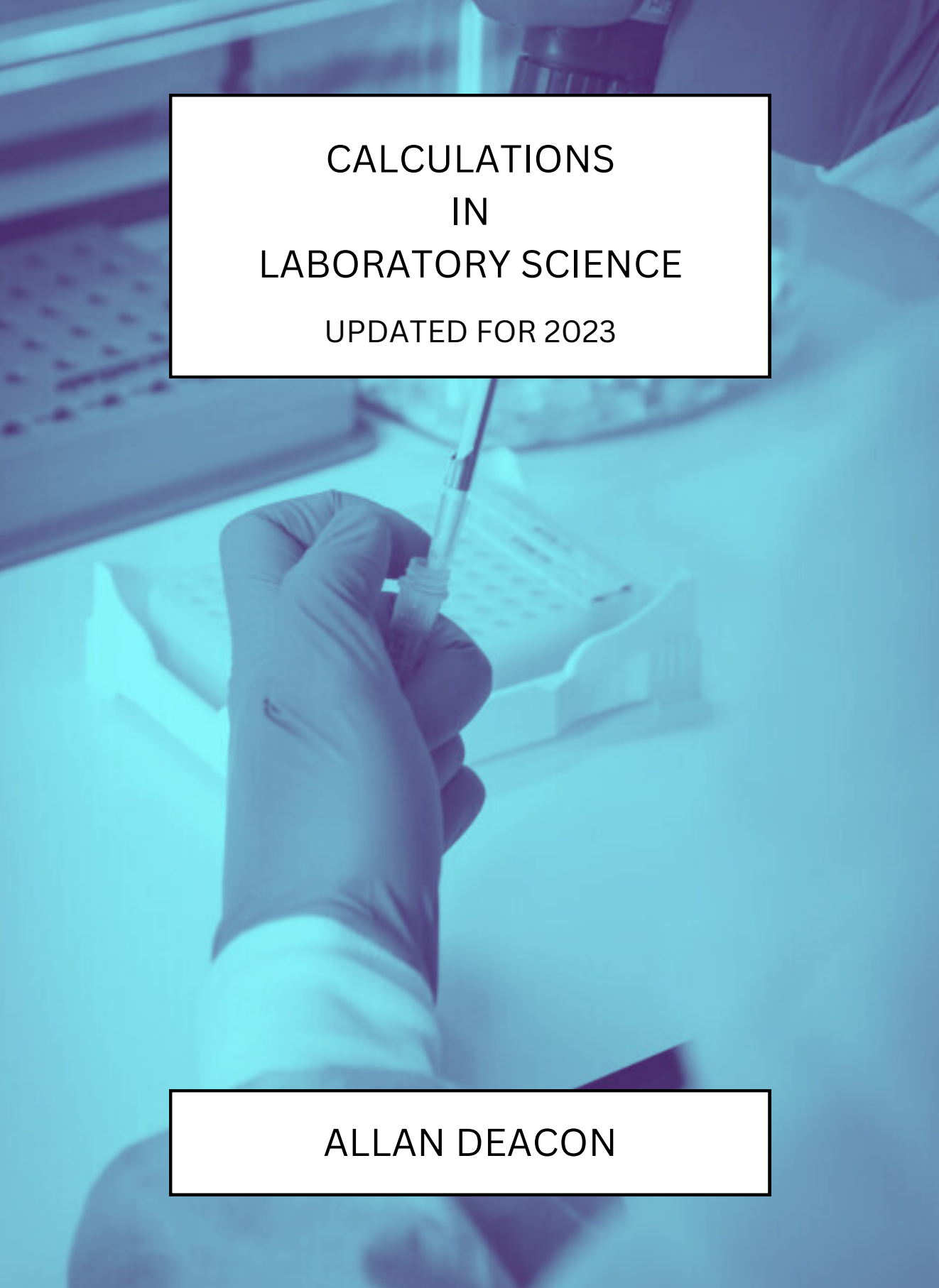


CALCULATIONS IN LABORATORY SCIENCE

UPDATED FOR 2023

ALLAN DEACON

CALCULATIONS IN LABORATORY SCIENCE

Allan Deacon BSc, PhD, FRCPath
Former Consultant Clinical Scientist, King's College and Bedford Hospitals NHS Foundation Trusts

Editors:
Roy Sherwood BSc, MSc, DPhil
Consultant Clinical Scientist, King's College Hospital, London

James Hooper MD, MSc, FRCPathConsultant Chemical Pathologist, Royal Brompton Hospital, London

ACB VENTURE PUBLICATIONS

DIRECTOR OF PUBLICATIONS AND COMMUNICATIONS
Chairman - Kamaljit Chatha

British Library Cataloguing in Publication Data

A catalogue record for the book is available from the British Library

ISBN 978-0-902429-43-7, EAN 9780902429437, ACB Venture Publications

Printed by Piggott Black Bear, Cambridge

Cover by Alan Sherwood, Aspire Design Studios, London

Preface

"What one fool can do, another can" (Ancient Simian Proverb)

Clinical biochemistry is predominantly a quantitative science in which both the acquisition and utilisation of laboratory data requires some understanding of the underlying mathematical principles involved. The Royal College of Pathologists acknowledges the importance of this by including calculation questions in the practical section of part 1 of its fellowship examination in chemical pathology (FRCPath). Examiners report that these questions are often done badly and candidates display not only poor numeracy skills, but a lack of understanding of the basic underlying physical chemistry and physiology. Sadly, calculations receive only scant attention in most textbooks and on most undergraduate and postgraduate courses.

I have previously tried to address this problem in two ways. Firstly, by holding tutorials: initially for local trainees in the Clinical Biochemistry Department at King's College Hospital and in later years on regional and national training courses organised by the Association for Clinical Biochemistry. Secondly, by publishing worked answers to past FRCPath examination calculation questions in the *ACB News*. However, trainees often express the need for a comprehensive textbook which not only presents worked examples, but brings together the relevant mathematics and basic science. This book is an attempt to meet that need.

This book was originally intended for trainees in clinical biochemistry, particularly those preparing for the FRCPath examination. However, I hope it will also prove useful to trainees in other clinical sciences and to undergraduate and postgraduate students in any of the life sciences. Each chapter takes a topic and explains the relevant physical chemistry and/or physiology. I have tried to derive the various formulae and mathematical procedures from first principles whenever possible in the belief that an understanding of their basis leads not only to their correct application, but helps the reader develop these methods to new problems which he/she may encounter in the future. Questions are included in the text, which I hope the reader will attempt before looking at the worked answer. At the end of each chapter there is a set of additional questions, the answers to which are provided in the Appendix.

Allan Deacon
March 2009

Contents

Important notice

Although ACB Venture Publications has made every effort to ensure the accuracy of the information contained in this book, the responsibility for the patient is ultimately that of the medical practitioner ordering or performing/supervising the investigations. All drugs and intravenous fluids must be prescribed by a registered medical practitioner and administered by an individual authorised to do so. The publishers, authors and editors do not assume any liability for any injury and/or damage to persons or property arising from this publication.

Chapter 1

Units and their manipulation

"A numerical result without units is meaningless"

Failure to pay attention to units is probably the commonest cause of error when performing even the simplest of calculations. In clinical biochemistry we often use several units for the same property. For example concentration may be expressed as mmol/L, mg/100 mL or g/L and volume may be expressed as L, mL or µL. Therefore ability to manipulate units is an essential prerequisite for the successful completion of most calculations.

Units used in clinical biochemistry

International standardization of units has obvious advantages. In the UK and Europe clinical laboratories have attempted to adopt *Systeme International d'Unites* (*SI* units), whereas in the USA mass units are still used.

The SI units for the three basic dimensions are:

Length - metre (m)
Mass - kilogram (kg)
Time - second (s)

Under this system, volume is defined as a cubic quantity of length, for example m^3, cm^3. However, most laboratories (and scientific journals) have retained the old metric unit of volume, the litre (L). 1L is equivalent to the SI volume of 1 dm^3.

Where the molecular weight (MW) of the substance being measured is known, the SI unit of quantity is the *mole* or multiple of a mole but the metric unit of volume, the litre, is still used. For example, glucose concentration is expressed as mmol/L. One *mole* is the formula weight of a substance measured in grams. Often the term molar is used instead of mol/L and is abbreviated to M. A term in common use for the symbol " / " is "per" and means the value obtained when one quantity is divided by another. Alternatively, the divisor may be written to the power of " -1". For example 1 mol/L may also be written 1 mol L^{-1}.

Question: Q1(1)

180 g of glucose (formula $C_6H_{12}O_6$) is dissolved in 1 L of distilled water. What is the concentration of glucose in (a) g/L, (b) mg/100 mL and (c) mol/L?

Answer: Q1(1)

a) Since 180 g of glucose was dissolved in 1L of water the concentration is **180 g/L.**

b) "Milli" means a thousandth of. Therefore 1 g of glucose contains 1000 mg.
Therefore 180 g contains 180 x 1000 = 180,000 mg.
Similarly, 1L of water contains 1000 mL which is 10 multiples of 100 mL.
Therefore the concentration is 180,000/10 = **18000 mg/100 mL.**
NB. mg/100 mL is sometimes written mg% or mg/dL.

c) The molecular weight of glucose is (6 x 12) + (12 x 1) + (6 x 16) = 180.
Therefore 180 g of glucose contains 180/180 = 1 mole. Since it is dissolved in 1 litre of water the concentration is **1 mol/L or 1 mol L^{-1}** or **1 M.**

When the molecular weight is not known, or the analyte being measured is a mixture of substances of differing molecular weights (such as plasma total protein) then mass units may be used. e.g. plasma total protein is expressed as g/L.

When concentration is measured indirectly by some property which is not easily related to a given mass of analyte (e.g. enzyme activity), or using a method calibrated against a standard of unknown purity (e.g. some hormones) then arbitrary units are used. e.g. creatine kinase activity is expressed as U/L. When an internationally accepted standard preparation is used to calibrate the assay, results may be expressed as IU/L.

Dealing with both large and small numbers

Units are often adapted by adding a prefix to simplify dealing with very small and very large numbers. For example, volume may be expressed in litres, mL or µL and concentration may be expressed as mmol/L or µmol/L, which may cause problems, for example, when calculating glomerular filtration rate from creatinine concentrations expressed as µmol/L in blood and mmol/L in urine, particularly when by convention the result is required in mL/min. This practice has arisen in order to avoid numbers which are clumsy. For example, 25 µL is far more convenient to handle than 0.000025 L. A list of common prefixes is given in Figure 1.1.

Name	Symbol	Meaning	Example
kilo-	k	10^{3} x	kg body weight
deci-	d	10^{-1} x	dL (0.1L or 100 mL)
centi-	c	10^{-2} x	cm (1/100th of a metre)
milli-	m	10^{-3} x	mmol (1/1000th of a mole)
micro-	µ	10^{-6} x	µL (1/1,000,000 th of a litre)
nano-	n	10^{-9} x	nm (1/1,000,000,000th of a metre)
pico-	p	10^{-12} x	pg (1/1,000,000,000,000th of a gram)
femto-	f	10^{-15} x	fmol (1/1,000,000,000,000,000 of a mole)

Figure 1.1 Prefixes for powers of 10 with common applications

An alternative approach is to express a numerical value as a number in the range 1-10 multiplied by an appropriate power of 10. For example, the weight 0.005 g could be written as 5.0×10^{-3} g as well as 5 mg. As will be discussed later, this practice enables a numerical result to be expressed to a given number of significant figures.

The general term 10^x means 10 multiplied by itself x number of times, so that the expression $y \times 10^x$ means y multiplied by 10, x times. Since our number system is based on 10 all this means is moving the decimal place to the right x number of times. For example 3.125×10^2 simply means that the decimal point is moved 2 places to the right to become 312.5. The number 14×10^3 does not have a decimal point (or to be correct it is implied that the decimal point is placed after the last digit, i.e. 14) so instead of moving the decimal point 3 digits to the right 3 noughts are added to give 14000. Similarly, for 6.25×10^4 the decimal point is moved two places to the right then two noughts added to give 62500.

On the other hand, 10^{-x} means 1 divided by 10^x i.e. $1/10^x$ (in a similar way that 1 mol L^{-1} means 1 mol/L). Therefore, instead of moving the decimal point one place to the right or adding a nought for each increment in x, the opposite applies and the decimal point is moved to the left; when the number of digits is exhausted, noughts are placed in front of it. For example 6834×10^{-3} means 6.834, 24.52×10^{-2} means 0.2452 and 6.35×10^{-4} means 0.000635.

Significant figures

The way in which a numerical value is written makes a statement about the reliability or accuracy of that value. For example, the concentration of an analyte determined by a colourimetric assay written as 0.103562 mol/L implies a greater degree of accuracy (and hence precision of its method of measurement) than if it were written 0.104 mol/L. This result must have been calculated from an absorbance reading with only three (or possibly four) digits, so it is misleading to quote the result to 6 significant figures. The reliability of the measurement process permits the reporting of the result to only 3 significant figures at the most.

As a general rule one should express a result with no more (or only a little more) precision than the accuracy of the data from which it was calculated. When the result is calculated from more than one piece of data then the accuracy of the least precise piece of data should be used.

For example, consider a creatinine clearance calculated from plasma and urine creatinine measurements and a 24 h urine volume of 1836 mL. Creatinine measurements are made to only 3 figures and the reliability of the third must be doubtful. Although the volume is known to 4 figures, the accuracy of timed urine collections is notoriously poor and the volume is only likely to be accurate to 2 figures at the most.

The accuracy of the final result of a calculation can never be greater than that of the least accurate measurement used in its calculation.

When the final result has been calculated and clearly has a greater number of significant figures than is warranted, the unwanted digits can be removed either by rounding or truncation. In rounding, the value of the retained digit is increased by 1 only if the discarded digit(s) begin with 5, 6, 7, 8 or 9. In truncation the unwanted digits are simply removed. For example the value 2.3478 can be rounded to 2.35 or truncated to 2.34. Rounding is slightly more accurate than truncation and is preferred.

During the intermediate stages of a calculation, errors can accumulate if intermediate values are rounded off to the desired final number of significant figures. A fairly reliable rule is:

Use one more significant figure in the calculations than one expects to retain in the final result.

The exception to this rule is calculations which involve small differences between nearly equal numbers. For example, if a beaker weighs 20.4675 g empty but 20.5796 g after addition of a sample then it would be wrong to subtract 20.5 from 20.6 since this would give only one significant figure in the answer, whereas the data would clearly support more.

Sometimes it is difficult to convey the number of significant figures in an answer. For example a result of 100 g contains only one significant figure (zeros do not count!) One solution is to place a decimal point after the result so that it becomes 100. which indicates three significant figures. It is more difficult to convey two significant figures. One solution it to change the units and express the result as 0.1 kg (1 significant figure), 0.10 kg (2 significant figures) or 0.100 kg (3 significant figures). Alternatively the result could be expressed as a number multiplied by 10 to the power of another number e.g. 1×10^2 g. It is then easy to convey the number of significant figures e.g. 1×10^2 (one significant figure), 1.0×10^2 (two significant figures) or 1.00×10^2 (3 significant figures) etc.

What happens to units during a calculation?

There are two general rules which must be remembered:

- For the operations of addition and subtraction, the dimensions and the units must be the same and remain unchanged after the calculation.

- For the operations of multiplication and division, the dimensions are multiplied and divided just as are the numbers, the result being the product or quotient of the dimensions.

For example, to add together the two weights 0.952 g and 0.23 mg it is necessary to either convert the first weight to mg then add it to the second, yielding a result in mg, or, to convert the second weight to grams then add it to the first, in which case the answer will be in g.

The calculation of molar absorptivity involves combining several units. The expression for calculating molar absorptivity (ε) is:

$$\varepsilon = \frac{\text{Absorbance}}{\text{Concentration} \times \text{path length}}$$

Where as absorbance does not have any units, concentration is in mol/L and the optical path length is expressed in cm. Substituting these units into this expression gives:

$$\varepsilon = \frac{1}{\text{mol/L} \times \text{cm}} = \frac{\text{L}}{\text{mol} \times \text{cm}}$$

so that the units for ε are L/mol/cm or L. mol^{-1} cm^{-1}.

Care needs to be taken when calculating ratios of concentrations. If the two analytes and their units are identical, then the result does not have any units. For example, to calculate the ratio of plasma to urinary calcium, when both are expressed in the same units, then the units cancel:

$$\frac{\text{Plasma calcium (mmol/L)}}{\text{Urine calcium (mmol/L)}} = \frac{\text{mmol} \times \text{L}}{\text{L} \times \text{mmol}}$$

If both concentrations are in mass units, then both concentrations can be converted to SI units by multiplying the mass concentration by the atomic weight (40). Since both operations are identical, the ratio will be the same and will not have any units:

$$\frac{\text{Plasma calcium (mmol/L)}}{\text{Urine calcium (mmol/L)}} = \frac{\text{Plasma calcium (mg/L)} \times 40}{\text{Urine calcium (mg/L)} \times 40}$$

If the concentrations are for different analytes then this is no longer true. The ratio of their concentrations calculated from mass units will be different to that calculated from SI units. For example, calculation of the calcium:creatinine ratio in a random urine in which the concentration of calcium is 2.5 mmol/L and that of creatinine is 5.0 mmol/L gives:

$$\frac{\text{Urine calcium}}{\text{Urine creatinine}} = \frac{2.5 \text{ mmol/L}}{5.0 \text{ mmol/L}} = 0.5$$

If the concentrations are converted to mass units by multiplication of the calcium concentration by its atomic weight (40) and the creatinine concentration by its molecular weight (113) then a different ratio is obtained:

$$\frac{\text{Urine calcium (mg/L)}}{\text{Urine creatinine (mg/L)}} = \frac{2.5 \text{ (mmol/L)} \times 40}{5.0 \text{ (mmol/L)} \times 113} = 0.18$$

Thus whilst SI concentrations of different analytes are comparable, mass concentrations are not. 1 mg of calcium is not equivalent to 1 mg of creatinine. To avoid confusion it is best to give the units of the components of the ratio in parentheses after the ratio i.e. 0.5 (mol/mol) or 0.18 (g/g).

Question Q1(2)

a) 45 mL of solution A and 2.65 L of solution B are mixed. What is the total volume?

b) How many grams of a substance are contained in 350 mL of a solution containing 50 g/L?

Answer Q1(2)

a) Total volume = Volume A + Volume B

The units of the two volumes are different (i.e. 45 mL and 2.65 L)
Before they can be added together one must be converted to the other so that the units are the same. There are 1000 mL in a L, therefore if 45 mL is divided by 1000 then it becomes 0.045L. (Alternatively 2.65 L could be multiplied by 1000 to convert it to 2650 mL).

Total volume = 0.045 L + 2.65 L = **2.695 L (or 2695 mL)**

b) Amount of substance = volume x concentration

Both volume and concentration contain a volume term, but their units are different i.e. mL and L. These need to be converted to the same units. If the volume is converted from mL to L by dividing by 1000 then it becomes 0.35 L So that:

Amount of substance = 0.35 x 50 = 17.5

To find out what units to use, carry out the calculation with units rather than numbers:

Amount of substance = L x g/L which can be written $\frac{L \times g}{L}$

Since the litres (the Ls) cancel, then the final units are g
so the amount of substance contained in 350 mL is **17.5 g.**

Interconversion of mass and SI units

Since SI units of concentration used in clinical biochemistry are mol/L or multiples of mol/L i.e. mmol/L, µmol/L etc, then conversion to and from mass units requires knowledge of the molecular or atomic weight. The two sets of units are related by the expression:

$$\textbf{Concentration (mol/L)} = \frac{\textbf{Concentration (g/L)}}{\textbf{Molecular or atomic weight}}$$

This relationship still applies to units with different prefixes as long as the same prefix is used on both sides of the equation. For example if the mass concentration is in mg/L then in SI units the concentration will be in mmol/L and not mol/L; and if the mass concentration is in mg/100 mL then the SI units will be mmol/100 mL. The above relationship can be manipulated to carry out the reverse calculation, e.g. to convert mmol/L to mg/L.

Use of equivalent weights

Chemists have, in the past, attempted to simplify calculations by introducing the concept of equivalent weight. The equivalent weight of one substance reacts exactly with the equivalent weight of another. For example in titrimetric analysis one mole of hydrochloric acid neutralises one mole of sodium hydroxide, whereas one mole of sulphuric acid (which yields two titratable hydrogens) will neutralize two moles of sodium hydroxide. Therefore one mole of sulphuric acid is equivalent to two moles of hydrochloric acid. These principles also apply to redox reactions involving metal ions. Monovalent ions, such as sodium and potassium, are equivalent to one hydrogen ion so that their equivalent weights are equal to their atomic weights. However, divalent ions , such as calcium and magnesium are equivalent to two hydrogen ions and their equivalent weight is one half their atomic weights. In general:

$$\text{Equivalent weight} = \frac{\text{molecular weight}}{\text{Valency}}$$

$$\text{number of equivalents} = \frac{\text{weight in g}}{\text{Equivalent weight}}$$

$$\text{number of equivalents} = \frac{\text{weight in g} \times \text{valency}}{\text{molecular weight}}$$

In the case of univalent ions the units will be numerically the same e.g. 140 mmol/L of sodium is the same as 140 mEq/L. For divalent ions 1 mol contains 2 Eq, e.g. 1 mmol/L of magnesium is the same as 2 mEq/L. Results are no longer reported as mEq/L in the UK but may be found in the literature. The term “normal,” often abbreviated as N, may also be encountered which is the concentration in equivalents per litre e.g. 1 molar (or M) saline (1 mol/L) is equivalent to 1 normal (or N) saline (1 Eq/L). This should not be confused with “normal” or “physiological saline (9 g/L).

Question: Q1(3)

a) Express 360 mg% glucose as mmol/L
b) Express 140 mmol/L saline as g/100 mL
c) Convert 5.64 mEq/L calcium to mmol/L.

Answer Q1(3)

a) $\text{mmol/L} = \frac{\text{mg/L}}{\text{MW}}$

First calculate the molecular weight of glucose – formula $C_6H_{12}O_6$

C_6	=	Atomic weight of glucose	x	6	=	12	x	6	=	72
H_{12}	=	Atomic weight of hydrogen	x	12	=	1	x	12	=	12
O_6	=	Atomic weight of oxygen	x	6	=	16	x	6	=	96
								SUM:		180

Glucose concentration is 360 mg/% i.e. 360 mg/100 mL

Concn (mg/L) = concn (mg/100mL) x 10 = 360 x 10 = 3600 mg/L

And glucose concn (mmol/L) = $\frac{3600}{180}$ = **20 mmol/L**

b) g/L = mol/L x MW

Sodium chloride concn = 140 mmol/L = $\frac{140}{1000}$ = 0.14 mol/L

(Since there are 1000 mmol in a mol)

Molecular weight of sodium chloride (NaCl) = 23 + 35.5 = 58.5

NaCl (g/L) = 0.14 x 58.5 = 8.19 g/L

And NaCl (g/100 mL) = $\frac{\text{g/L}}{10}$ = $\frac{8.19}{10}$ = **0.82 g/100 mL** (2 sig figs)

d) Calcium is a divalent cation (Ca^{++}) and therefore equivalent to two hydrogens

Therefore mEq/L = mmol/L x 2

And calcium concentration (mmol/L) = $\frac{\text{mEq/L}}{2}$ = $\frac{5.64}{2}$ = 2.82 mmol/L

Blood urea nitrogen (BUN)

It is common practice in the USA to express plasma urea concentration in terms of its contribution to the nitrogen content of plasma, usually using mg% (i.e. mg/100 mL or mg/dL) as units. The formula of urea is $CO(NH_2)_2$ so that each molecule contains 2 nitrogen atoms. Therefore one mole of urea contains one mole of molecular nitrogen (N_2) and the molecular weight of nitrogen (N_2) is twice the atomic weight (14) i.e. 28. Since:

$$\text{Concentration (g/L)} = \text{Concentration (mol/L)} \times \text{Molecular weight}$$

we can write for urea:

$$\text{BUN (mg/L)} = \text{Urea (mmol/L)} \times 28$$

and if BUN is to be expressed as mg/100 mL (mg%), then both sides are divided by 10:

$$\text{BUN (mg/100mL)} = \frac{\text{Urea (mmol/L)} \times 28}{10} = \text{Urea (mmol/L)} \times 2.8$$

Alternatively, working in atoms of nitrogen rather than moles of molecular nitrogen, the atomic weight of 14 is used but the molar concentration of urea is still multiplied by 2 since one mole of urea contains 2 atoms of nitrogen (as compared to one mole of nitrogen):

$$\text{BUN (mg/L)} = 2 \times \text{urea concentration (mmol/L)} \times \text{atomic weight (14)}$$

The overall result is the same. The decision of whether to work in atoms or moles of nitrogen often causes confusion. The two expressions for the inter-conversion of BUN and urea concentrations are:

$$\text{BUN (mg/100mL)} = \text{Urea (mmol/L)} \times 2.8$$

$$\text{Urea (mmol/L)} = \frac{\text{BUN (mg/100mL)}}{2.8}$$

Question Q1 (4)

Express BUN 14 mg% as urea concentration in mmol/L.

Answer Q1 (4)

$$\text{Urea (mmol/L)} = \frac{\text{BUN (mg\%)}}{2.8}$$

Substituting BUN = 14 mg%

$$\text{Urea (mmol/L)} = \frac{14}{2.8} = \mathbf{5\ mmol/L}$$

Further questions

Atomic weights: C = 12; H = 1; O = 16; Ca = 40; N = 14

1. Convert the following: a) 125 mg% to g/L; b) 0.25 mol/L to mmol/L; c) 0.236 nmol/L to µmol/L; d) 1.6 mg/L to ng/mL.

2. Convert the following concentrations to "SI" units: a) plasma glucose 120 mg%; b) serum calcium 4.0 mEq/L; c) BUN 21 mg%; d) Serum creatinine 0.66 mg%.

3. Convert the following: a) plasma glucose from 20 mmol/L to mg/100 mL; b) serum calcium from 3.2 mmol/L to mEq/L; c) serum urea from 30.6 mmol/L to mg% BUN; d) serum creatinine from 250 µmol/L to mg%.

4. Convert the following: a) 1.5×10^{-3} M to mmol/L; b) 1.25×10^{-5} M to µmol/L; c) 2.5×10^{2} mg/100 mL to g/L; d) 3.25×10^{-6} mmol/L to µmol/L.

5. After incubation of an enzyme with substrate for 30 min the concentration of product in the reaction mixture was 3.00×10^{-3} M. a) How many mmol of product would be present in 100 mL of the reaction mixture; and b) what is the rate of formation of product in 250 mL of reaction mixture expressed as µmol/min?

6. If an acid dissociates in solution to give its conjugate base and hydrogen ions, what are the units of its dissociation constant if urine contains 0.1 M of undissociated acid, 25×10^{-5} mol/L of its conjugate base and 120 nmol/L of hydrogen ions? NB the dissociation constant is the product of the concentrations of conjugate base and hydrogen ions divided by the concentration of undissociated acid.

Chapter 2

Laboratory manipulations

In this chapter calculations involved in common laboratory manipulations such as preparing solutions of desired concentrations, preparing dilutions from a stock solution, calculating concentrations obtained on mixing solutions etc are described.

Preparing solutions from a solid material

If the required concentration is in mass units then calculation of the amount to be weighed out is a relatively simple matter provided attention is paid to the units involved For example to prepare 1 L of a solution containing 10 g/L, 10 g of the substance will need to be weighed out. The volume of solution required needs to be taken into consideration as well as the final concentration. To prepare 500 mL of the same solution then half the amount would be required i.e. 5 g. In general:

Weight required = Required concentration x volume

It is important that the units should be compatible. If the concentration is in g/L then the volume should be expressed as litres and the calculated amount to be weighed out will be in grams.

Question Q2(1)

Calculate the number of grams of glucose needed to prepare 2 L of a solution with a concentration of 150 mg/dL (150 mg%).

Answer Q2(1)

Since the weight is required in grams and the final volume in litres then the final concentration is first converted from mg/dL to g/L.

Since there are 1000 mg in a gram, 150 mg is equivalent to 150 / 1000 = 0.15 g

Since there are 10 dL in 1 L then 1 dL is equivalent to one tenth of a litre = 0.1 L

Therefore the final concentration is 0.15 g / 0.1 L. Multiplication by 10 converts to g/L i.e. 0.15 x 10 = 1.5 g/L

Therefore weight required (g) = Concentration (1.5 g/L) x final vol (2 L) = 3.0 g

---

If the target concentration is given in SI units then the weight of the substance to be weighed out must be calculated using both the molecular weight (MW) of the substance and the final volume required. It is usually simplest to first convert the target concentration from SI to mass units:

Concentration (mass units) = Concentration (SI units) x Molecular weight

Again the prefix to the concentration terms (i.e. m, n or μ) must be the same for both the mass and SI units. For example, if the SI concentration is in mmol/L then the mass concentration will be in mg/L.

Suppose we needed to prepare 500 mL of a solution of sodium chloride with a concentration of 140 mmol/L. The first thing would be to calculate the concentration of sodium chloride in mass units (i.e. mg/L):

The atomic weights of sodium and chlorine are 23 and 35.5 respectively. Therefore the molecular weight of sodium chloride (NaCl) is 23 + 35.5 = 58.5.

NaCl concentration (mg/L) = 140 (mmol/L) x MW (58.5) = 8190 mg/L

The amount of NaCl required to prepare 500 mL of solution will be half of this: 8190 / 2 = 4095 mg. Since most balances have scales in g rather than mg, division by 1000 (since there are 1000 mg in a g) gives the weight in g (4095/1000 = 4.095 g).

Sometimes the chemical required to prepare a solution may not be in exactly the same form as that described in a method. For example, a method sheet for a manual glucose

method dictates that 1 L of a 500 mg% solution of glucose is prepared by dissolving 5.00 g of glucose in water and making the final volume up to 1 L. However, if only glucose monohydrate is available then a glucose solution containing the same concentration of glucose can still be prepared if the difference in molecular weights is taken into account. First convert the glucose concentration to SI units. The molecular weight of glucose (formula $C_6H_{12}O_6$) is 180. Since 5.00 g of glucose is normally weighed out and made up to 1L the final concentration is 5.00 g/L.

$$\text{Therefore glucose (mol/L)} = \frac{\text{glucose (g/L)}}{\text{Molecular weight}} = \frac{5.00}{180} = 0.0278 \text{ mol/L}$$

Since one mole of glucose monohydrate (formula $C_6H_{12}O_6.H_2O$) contains one mole of glucose, then the concentration of glucose monohydrate will also be 0.0278 mol/L. Conversion of the glucose monohydrate concentration to mass units will give the weight of glucose monohydrate to be weighed out.

MW of glucose monohydrate = MW glucose + MW water = 180 + 18 = 198

Glucose monohydrate (g/L) = Glucose monohydrate (mol/L) x MW

= 0.0278 x 198 = **5.50 g/L**

Therefore 5.50 g of glucose monohydrate will need to be weighed out instead of 5.00 g of glucose. Note that the same result can be obtained by multiplying the weight of glucose by the ratio of the molecular weights of the hydrated to the anhydrous form:

$$\text{Wt glucose monohydrate (g)} = \text{wt glucose (5.00g)} \times \frac{198}{180} = \mathbf{5.50\ g}$$

Question 2(2)

a) How many grams of anhydrous disodium hydrogen phosphate will be needed to prepare 2 litres with a concentration of 50 mmol/L.

b) Instructions for preparing 1L of a phosphate buffer state that 12.00 g of anhydrous sodium dihydrogen phosphate are required. If this material is unavailable how many grams of sodium dihydrogen phosphate dihydrate would be required?

(MWs: Na = 23, P = 31)

Answer Q2(2)

a) First calculate MW of anhydrous disodium hydrogen phosphate (Na_2HPO_4):

Na_2	=	2 x 23	=	46
H	=	1	=	1
P	=	31	=	31
O_4	=	4 x 16	=	64
				142

Next convert the concentration from mmol/L to mol/L then to g/L:

Since there are 1000 mmol in 1 mol,

50 mmol/L is the same as $\frac{50}{1000}$ = 0.05 mol/L

Concentration (g/L) = Concentration (mol/L) x MW

= 0.05 x 142 = 7.1 g/L

To prepare 2 L, twice this amount will be needed, i.e. 2 x 7.1 = **14.2 g**

b) MW NaH_2PO_4 = 23 + (2 x 1) + 31 + (4 x 16) = 120

MW $NaH_2PO_4.2H_2O$ = 120 + 2 (2 + 16) = 156

12.0 g/L of NaH_2PO_4 is equivalent to $\frac{12.0}{120}$ = 0.1 mol/L

0.1 mol/L $NaH_2PO_4.2H_2O$ contains $\frac{12.0 \times 156}{120}$ = **15.6 g/L**

Correcting for purity

For many chemicals used in the laboratory the percentage purity is significantly less than 100%. If a compound has a purity of *x* %, this means that each 100 g of the material contains *x* g of the compound. It follows that each g will contain only *x*/100 g. Therefore, to prepare a solution containing W g/L then the weight required is W x 100 /*x* g/L. In general:

$$\text{Weight of impure material} = \frac{\text{Weight of pure material} \times 100}{\%\ \text{purity}}$$

Again the weight units must be the same on both sides of the expression.

Preparing solutions from liquids

Some chemicals used to prepare reagents are themselves liquids. If the liquid is weighed then the procedure is the same as for solids, providing allowance is made for any departure from a purity of 100%. However, if the liquid is measured in volume then allowance must be made for the fact that most liquid chemicals do not have a density of one.

The units of density are weight/volume. If a liquid has a density of *x* g/mL, then this means that each mL contains *x* g of the compound. In general:

$$\text{Density} = \frac{\text{weight}}{\text{volume}}$$

The term specific gravity (SG) is often used. This is the density expressed as a ratio to the density of water (which is 1 g/mL). For most practical purposes SG and density can be considered as being the same thing (although SG, being a ratio, does not have units). By re-arranging the above equation, it is a simple matter to calculate the volume which contains the target weight of a compound. Suppose we wished to prepare 1 L of an ethanol standard solution, containing 800 mg/L of ethanol, from ethanol which has an SG of 0.79. The units for density must be the same as for concentration, the density (=SG) is in g/mL, the weight of ethanol per litre must also be in g (i.e. 800 mg / 1000 = 0.8 g /L).

$$\text{Volume (mL)} = \frac{\text{weight (g)}}{\text{Density (g/mL)}} = \frac{0.8}{0.79} = 1.01\ \text{mL}$$

Question Q2(3)

How many mL of hydrochloric acid (SG 1.16) are required to prepare 500 mL of 2.0 molar hydrochloric acid. The purity of the acid is 32 % w/w.

Answer Q2(3)

MW hydrochloric acid (HCl) = 1 + 35.5 = 36.5

Weight (g) of pure acid required to make 1 L 2.0 M HCl = 2.0 x 36.5 = 73 g

Weight required to make 500 mL 2.0 M HCl = 73 / 2 = 36.5 g

Since HCl has a purity of 32 %w/w, the weight of HCl (SG 1.16) required is

$$\frac{36.5 \times 100}{\%\ \text{purity}} = \frac{36.5 \times 100}{32} = 114.0\ \text{g}$$

Using the density of 1.16, the volume can be calculated:

$$\text{Volume (mL)} = \frac{\text{weight (g)}}{\text{Density (g/mL)}} = \frac{114.0}{1.16} = \mathbf{98.3\ mL}$$

Dealing with dilutions

In the laboratory we often need to calculate the final concentration of a substance after a given dilution, the volume of a stock solution which has to be diluted to give a target final concentration or how much liquid to add to a set volume of a stock solution to give a desired concentration. All of these problems are variations on a single theme and are approached in the same way. First it is important to realise that the total amount of a substance (whether expressed in mass or SI units) in a solution is the product of concentration and volume (the volume in the concentration term must be in the same units as the volume of solution).

Total amount = concentration x volume

For example:

2 L of 0.1 M sodium hydroxide (0.1 mol/L) will contain 2 x 0.1 = 0.2 mols

500 mL of 100 mM glucose (100 mmol/L) will contain 0.5 x 100 = 50 mmol

1.5 L of 20% sodium chloride (20% = 20 g/100 mL = 200 g/L) contains 1.5 x 200 = 300 g

If a finite amount of solution is diluted with solvent then the total amount of the solute in the final solution will be the same. It is only the volume (which has become larger) and the concentration (which has become lower) that have changed:

Initial amount of solute = Final amount of solute

Since the amount of a solute is the product of concentration and volume, then the following expression can be written:

initial concentration x initial volume = final concentration x final volume

All dilution problems can be solved using this equation. If any three of the terms are known then this formula can be rearranged to obtain the remaining term. In the laboratory there are four applications which are commonly used:

1. The concentration of an analyte in a biological sample may exceed the working range of the assay. In this situation it is common practice to dilute the sample before carrying out the analysis and then to calculate the concentration in the undiluted sample. For example, 0.1 mL of urine is diluted to 2.0 mL with water; the creatinine concentration measured on the diluted urine is 150 µmol/L.

Initial concentration	= ?	Final concentration	= 150 µmol/L
Initial volume	= 0.1 mL	Final volume	= 2.0 mL

initial concentration x initial volume = final concentration x final volume

This equation can be rearranged if both sides are divided by the initial volume (the initial volume terms on the left hand side cancel each other).

$$\text{initial concentration} = \frac{\text{final concentration} \times \text{final volume}}{\text{initial volume}}$$

$$= \frac{150 \times 2.0}{0.1} = 3000 \text{ µmol/L} = \mathbf{3.0\ mmol/L}$$

In other words the result for the diluted sample is simply multiplied by the dilution (2.0/0.1 = 20). If the dilution is prepared manually, it is often simpler to add 0.1 mL of urine to 2.0 mL of water, giving a final volume of 2.1 mL. The method of calculation is exactly the same, and the final dilution will be 21 (2.1/0.1) instead of 20.

2. Calculation of the volume of a stock solution of a chemical which will need to be diluted to produce a given volume of solution with a target concentration. In this situation the initial volume is unknown but the initial concentration, final volume and final concentration are all known. For example, we may wish to calculate how much 1.0 M hydrochloric acid will need to be diluted to 100 mL to give a concentration of 0.025 M.

Initial concentration = 1.0 M Final concentration = 0.025 M
Initial volume = ? Final volume = 100 mL

$$\text{initial volume} = \frac{\text{final concentration} \times \text{final volume}}{\text{Initial concentration}}$$

$$= \frac{0.025 \times 100}{1.0} = 2.5 \text{ mL}$$

3. Calculation of the volume of diluent to be added to a given volume of stock solution to achieve a target concentration. This is often necessary if we wish to make best use of all of a stock solution remaining in a near empty bottle. For example, if 325 mL of a stock diluent containing 0.5 M phosphate buffer remains, how much water will need to be added if we wish to prepare the maximum volume of buffer with a phosphate concentration of 0.05 M.

Initial concentration = 0.5 M Final concentration = 0.05 M
Initial volume = 360 mL Final volume = ?

$$\text{Final volume} = \frac{\text{Initial concentration} \times \text{initial volume}}{\text{Final concentration}}$$

$$\text{Final volume} = \frac{0.5 \times 360}{0.05} = 3600 \text{ mL}$$

The volume to be added can be calculated from the initial and final volumes:

Final volume = initial volume + volume added

Volume added = final volume - initial volume

= 3600 - 360 = **3240 mL**

4. Calculation of the final concentration achieved when a known dilution is prepared from a solution of known initial concentration. For example, if 0.1 mL of a standard solution containing 50 mmol/L glucose is added to 2.0 mL of diluent, what is the final concentration?

Initial concentration = 50 mmol/L Final concentration = ?
Initial volume = 0.1 mL Final volume = 0.1 + 2.0 = 2.1 mL

$$\text{Final concentration} = \frac{\text{Initial concentration x initial volume}}{\text{Final volume}}$$

$$= \frac{50 \times 0.1}{2.1} = \textbf{2.38 mmol/L}$$

Question Q2(4)

A *working* reagent for a phosphate assay is prepared by mixing 100 mL of *stock* reagent with 900 mL of diluent. If only 360 mL of diluent is available, how much *stock* reagent must be added to obtain the maximum volume of *working* reagent?

Answer Q2(4)

In this situation, the initial concentration of the *stock* reagent can be regarded as 100%. The final concentration in the diluted stock will be 10% (since 1 part of *stock* reagent is mixed with 9 parts of *diluent* giving a *working* dilution of 1 in 10).

Stock reagent	*Working* reagent
Initial concentration = 100%	Final concentration = 10%
Initial volume = Volume *stock* mL	Final volume = (360 + Volume *stock*) mL

Volume *stock* (mL) x 100% = (360 + Volume *stock*) mL x 10%

Expanding brackets on the right hand side of the equation:

Volume *stock* x 100 = 3600 + (Volume *stock* x 10)

Moving (volume *stock* x 10) to the left hand side:

(Volume *stock* x 100) - (Volume *stock* x 10) = 3600

Taking the "volume *stock*" term outside of the brackets, re-arranging and solving:

Volume *stock* (100 - 10) = 3600

$$\text{Volume } stock = \frac{3600}{(100 - 10)} = \frac{3600}{90} = \mathbf{40\ mL}$$

Preparing a series of dilutions

Often a set of concentrations of a substance will need to be prepared, either to construct a standard curve, investigate the effect of varying a constituent in an assay or to determine the K_m or K_i of an enzyme. The golden rule is to first prepare a suitable amount of the solution with the highest concentration, then to set about preparing dilutions from this. Two approaches can be used, depending on whether or not the steps in concentration need to be equal.

For example, to set up a manual glucose method then a likely set of standards to cover the range of clinical interest would be solutions containing 5, 10, 15, 20 and 25 mmol/L of glucose. The first step is to prepare a stock glucose solution containing the highest concentration of glucose required i.e. 25 mmol/L. The lowest concentration required is 5 mmol/L which is 1/5th the concentration of the stock standard. As the steps are equal (5 mmol/L difference between each adjacent standard) then the second lowest concentration would be 2/5th of the stock, the next 3/5th etc. If 1 mL of each standard is required, then the lowest standard will contain 1/5 of 1 which is 0.2 mL of stock, the next standard will require twice this (0.4 mL of stock) etc. To ensure that the total volume is the same (1 mL) the volume of diluent (water) will *decrease* by 0.2 mL each time the volume of stock increases by 0.2 mL:

Required concentration (mmol/L):	5	10	15	20
Volume of stock glucose (25 mmol/L)	0.2	0.4	0.6	0.8
Volume diluent (water)	0.8	0.6	0.4	0.2

The scale can be altered to either increase or decrease the volumes of standards prepared – as long as the ratios of stock standard to diluent remain the same. Intermediate concentrations can be introduced by extrapolation. For example to prepare a glucose standard containing 7.5 mmol/L, 0.3 mL of stock would be mixed with 0.7 mL of diluent, and to prepare a standard containing 2.5 mmol/L, 0.1 mL of stock would be mixed with 0.9 mL of diluent.

An alternative approach is to prepare doubling dilutions. This results in a series in which each increment doubles the concentration of the solute i.e. the increments are not equal. Doubling dilutions are extremely easy to prepare and afford a means of quickly spanning a wide concentration range if there is uncertainty of the working concentration range required. As before, a stock solution is prepared containing the highest concentration required. A series of tubes is prepared. Each tube contains either stock solution or diluent, the volume of which is equal to the final volume required for each dilution. The first tube contains stock solution only, all of the remaining tubes contain diluent only. An equal volume of diluent is added to the first tube and mixed. The concentration of the contents of the first tube is therefore ½ of the stock concentration. The same volume of the contents of the first tube are transferred to the second tube and mixed. The concentration in the second tube is therefore ½ of the concentration in the first tube (½ x ½ = ¼). The same volume of the contents of the second tube are then transferred to the third tube and mixed which gives a concentration ½ of the concentration in the second tube (½ x ¼ = 1/8). This process can be repeated *ad infinitum*.

For example, if a stock solution has a concentration of 200 mmol/L, then preparing serial dilutions would give the following concentrations:

Tube number	1	2	3	4	5	6	7
Vol stock solution (mL)	1.0	-	-	-	-	-	-
Vol diluent (mL)	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Final concentration (mmol/L)	100	50	25	12.5	6.25	3.125	1.563
1 / final concentration (L/mmol)	0.01	0.02	0.04	0.08	0.16	0.32	0.64

The main drawback (apart from the concentrations not being evenly spaced) is that awkward numbers are soon encountered. This need not be a problem if the data is subsequently processed and plotted by computer. However, if the reciprocal of concentration is to be plotted (e.g. to determine the K_m of an enzyme) then the numbers produced are much easier to handle. If the material is in short supply then preparing doubling dilutions affords an economical and easy way to produce a wide range of concentrations provided minimal volumes are used. Serological titrations always employ doubling dilutions.

Further questions

(Atomic weights: H = 1; C = 12; O = 16; P = 31; Na = 23; K = 39; Ca = 40; S = 32)

1. How many grams of albumin are required to prepare 100 mL of a solution containing 70 g/L?

2. Calculate the concentration of sodium ions (in mmol/L) in a solution prepared by dissolving 85 g of sodium chloride in 1 litre of water.

3. What weight of calcium carbonate must be dissolved in 500 mL of dilute acid to provide a calcium standard containing 5.0 mmol/L ?

4. A solution contains 5 % sucrose. How much of this solution would you dilute to prepare 500 mL of 1 % sucrose?

5. 50 µL of urine is added to 5 mL of water. What is the resulting dilution of the urine?

6. Concentrated sulphuric acid (SG 1.84) is 96% by weight H_2SO_4. Calculate the volume of concentrated acid required to prepare 1 L of 0.1M H_2SO_4.

7. The following solutions were mixed together:

   50 mL potassium chloride (5.0 g/L)
   100 mL sodium chloride (50 g/L)

   Calculate the molar concentrations of potassium, sodium, and chloride ions.

8. If you have available 650 mL of 95 % ethanol, how much water would you add to obtain the maximum volume of 65 % ethanol?

9. In order to prepare 1 L of a stock standard solution containing 0.2 mol/L, the appropriate amount of sodium dihydrogen orthophosphate dihydrate should be weighed out. Due to an error, the same weight of anhydrous sodium dihydrogen orthophosphate used. Working standard was prepared by taking 5 mL of this stock standard and diluting it to 250 mL. What is the phosphate concentration (in mmol/L) of the working standard?

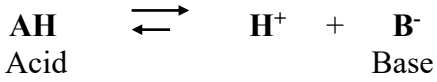
10. Solution A contains 12.0 g of anhydrous sodium dihydrogen phosphate per litre. What is the phosphate concentration expressed as mmol/L? What volume of solution A needs to be diluted to 1 L to give a phosphate concentration of 4 mmol/L.

Chapter 3

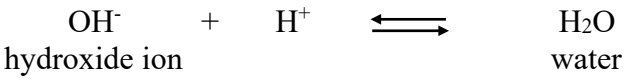
Acid-base, pH and buffers

What are acids and bases?

An acid is defined as a substance with a tendency to lose a proton (hydrogen ion) and a base as a substance with a tendency to gain a proton. It follows that there must be a relationship between an acid and a base. Whenever an acid loses a proton the anion formed will have a tendency to regain the proton, and hence it will be a base. Therefore, in general:

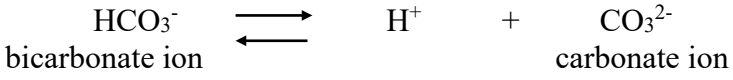


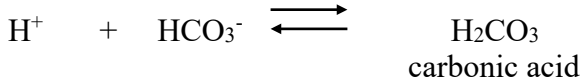
The acid and base (which differ only by a proton) are said to form a *conjugate pair*; every acid must have its conjugate base, and every base its conjugate acid e.g.



Strictly speaking, an alkali is any substance which can produce hydroxyl ions in aqueous solution but the terms base and alkali are often interchanged.

Some species, such as the bicarbonate ion can act as both proton donors and acceptors and therefore function as both an acid and a base:





Water can dissociate to produce a proton and its conjugate base (hydroxyl ion):

When as acid is added to water this equilibrium is driven to the left and the concentration of hydroxide ions decrease i.e. $H^+ >> OH^-$. When a base (e.g. hydroxide) is added to water then this equilibrium is again driven to the left but this time protons are removed i.e. $OH^- >> H^+$.

The fact that in aqueous solutions hydrogen ions are hydrated to form *hydroxonium* ions (H_3O^+) and free hydrogen ions do not exist is usually ignored.

How is the degree of acidity/alkalinity expressed?

Since most solutions we encounter are aqueous, water is taken as a reference point, and any solution in which the hydrogen ion concentration is greater than pure water is said to be acidic; any solution in which it is less is said to be alkaline. The *dissociation constant* (K) for water can be defined as:

$$K = \frac{[H^+][OH^-]}{[H_2O]} \quad \text{.....Eq.3.1}$$

where the brackets denote the concentrations of each species in mol/L. The concentration of water can be regarded as constant (since its concentration is very high and the molecule is only weakly ionized) and therefore the water concentration term is incorporated into the ionization constant, which is then referred to as the i*onic product of water* (K_w):

$$K_w = [H^+][OH^-] = 10^{-14} \text{ mol/L} \quad \text{..... Eq.3.2}$$

Since $[H^+] = [OH^-]$ then the concentration of each of these ions must be 10^{-7} mol/L. Therefore any solution in which $[H^+] > 10^{-7}$ mol/L will be *acidic* with the degree of acidity related to $[H^+]$. Conversely, any solution with $[H^+] < 10^{-7}$ mol/L will be *alkaline*.

The range of $[H^+]$ usually encountered by the chemist is very wide and quite often the $[H^+]$ is very low e.g. the $[H^+]$ of blood is typically 0.000000040 mol/L. In order to compress the scale and simplify the expression of low concentrations (and hopefully make calculations simpler) Sorensen, in 1909, devised the logarithmic pH scale.

What are logarithms?

Numbers can be written as some other number (called the *base*) raised to the power of another number (called the *logarithm*) i.e. $base^{logarithm}$. The power or logarithm defines the number of times the base value must be multiplied by itself to give the number. A few examples are given in Fig 3.1.

Mathematicians have found that working with logarithms has advantages in some situations. In clinical biochemistry we only need to use logarithms to two different bases: 10 (known as *common* logarithms) and 2.718 (known as *natural or Napierian* logarithms). Values for both types of logarithms are readily available on most pocket calculators or can be obtained from tables of logarithms. For example the logarithm of 31.62 is 1.5, a result which is difficult to obtain by manual calculation. In this chapter we will only use common logarithms (natural logarithms may be used in later chapters).

Number	Base	Logarithm	Evaluation
4	2	2	$4 = 2^2 = 2 \times 2$
8	2	3	$8 = 2^3 = 2 \times 2 \times 2$
16	2	4	$16 = 2^4 = 2 \times 2 \times 2 \times 2$
9	3	2	$9 = 3^2 = 3 \times 3$
16	4	2	$16 = 4^2 = 4 \times 4$
25	5	2	$25 = 5^2 = 5 \times 5$
125	5	3	$125 = 5^3 = 5 \times 5 \times 5$
100	10	2	$100 = 10^2 = 10 \times 10$
1000	10	3	$1000 = 10^3 = 10 \times 10 \times 10$

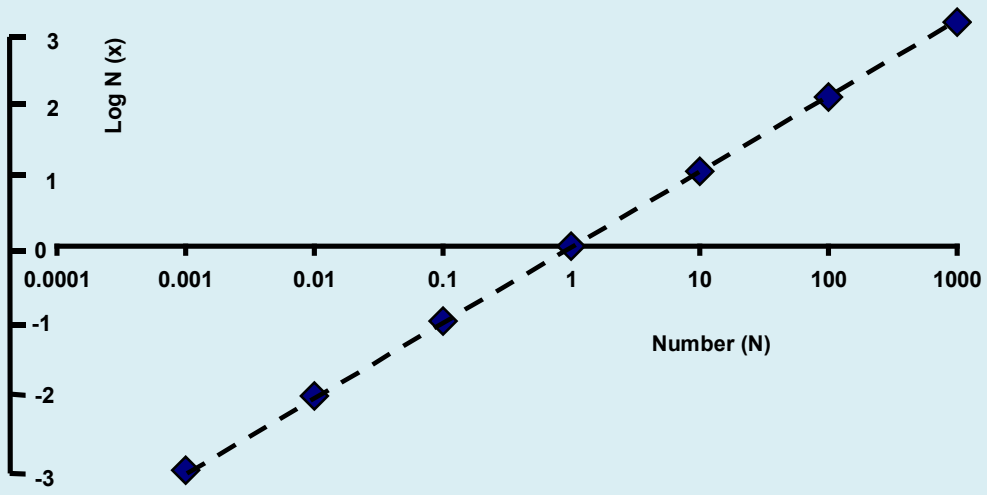
Figure 3.1 Examples of logarithms

The reverse of a logarithm is the *antilogarithm* which is the number which gave rise to the logarithm in the first place. The practical importance of antilogarithms is that at the end of a calculation we often end up with a result which is a logarithm; we then need to determine the number which would give rise to this antilogarithm. Fortunately antilogarithms are also available on most pocket calculators or can be obtained from tables of antilogarithms (or by using tables of logarithms backwards). The abbreviations "*log*" and "*antilog*" are respectively used for logarithm and antilogarithm.

A general notation relating a number (N) to its logarithm (x) to a base (b) can be used:

If, $N = b^x$		e.g. $100 = 10^2$
	then	
$\log_b N = x$		e.g. $\log_{10}100 = 2$
	and	
$\text{antilog}_b x = N$		e.g $\text{antilog}_{10}2 = 100$

Number (N):	**0.001**	**0.01**	**0.1**	**1**	**10**	**100**	**1000**
Log_{10}N (x):	**-3**	**-2**	**-1**	**0**	**1**	**2**	**3**



- **The logarithm of 1 is zero**
- **Numbers between 1 and zero have negative logarithms**
- **Logarithms do not have units**
- **Negative numbers do not have logarithms**

Figure 3.2 Relationship between numbers and their common logarithms

As illustrated in fig 3.2, $\log_{10}100 = 2$ and $\log_{10}1000 = 3$. This process can be continued indefinitely. For example, $\log_{10}10000 = 4$, i.e. $10000 = 10 \times 10 \times 10 \times 10 = 10^4$. Therefore, for common logarithms (where the base 10 is used), as a number increases by a factor of 10, it's log increases by 1. For example, increasing 100 by a factor of 10 gives 1000 whereas its log increases by 1, from 2 to 3. The converse is true, as a number decreases by a factor of 10 its log decreases by 1. What would happen if 100 was decreased by a factor of 10 to become 10? The log would have to decrease by 1 i.e. change from 2 to 1 so that $\log_{10}10 = 1$. This is not really surprising since the logarithm is the number of base terms which must be multiplied together to give the number; in this case the log is 1 which means that only one term is used so that the number is unchanged i.e. $10^1 = 10$. If 10 is decreased by a factor of 10 to give 1, then its log must also decrease by 1 and therefore becomes zero. This gives the surprising result that $\log_{10}1 = 0$. If this process is taken one stage further and 1 is decreased by a factor of 10 to 0.1 then its log decreases by 1 from 0 to -1, so that $\log_{10}0.1 = -1$. For numbers below 1 the logarithm becomes increasingly negative. As a consequence negative numbers cannot have logs and the log of zero has no meaning. Fig 3.2 illustrates the relationship between a number and its common logarithm.

What is pH? Why is it used?

pH is a logarithmic scale devised by Sorensen to simplify expression of the wide ranges in hydrogen ion concentration encountered by chemists. Since in most biological systems the hydrogen ion concentration is very low (much less than 1 mol/L) this inevitably means that the logarithm of the molar hydrogen ion concentration will be a negative number. For example, at a hydrogen ion concentration of 0.0000001 mol/L the logarithm would be -7. Negative numbers are inconvenient so the sign is changed to a positive value. *Therefore, a value can be defined called the pH which is* ***the negative logarithm to the base 10 of the molar hydrogen ion concentration.***

$$pH = -\log_{10}[H^+] \quad OR \quad pH = \log_{10}\frac{1}{[H^+]} \quad \text{..... Eq 3.3.}$$

Properties of pH are shown in Fig 3.3.

Question: Q3(1)

a) Calculate the hydrogen ion concentration of blood with a pH of 6.95.

b) Treatment with bicarbonate halves the hydrogen ion concentration. What is the new pH?

Answer to Q3(1)

a) Reverse the expression for pH so that the concentration term is on the left:

$$- \log_{10} [H^+] = pH$$

Multiply both sides by minus 1 in order to change the signs:

$\log_{10} [H^+] = -pH$, then take antilogarithms of both sides:

$$[H^+] = \text{antilog}_{10} (- pH)$$

substitute pH = 6.95 and evaluate:

$$[H^+] = \text{antilog}_{10} (- 6.95) = 0.000000112 = 1.12 \times 10^{-7} \text{ mol/L}$$

Multiply by 10^9 to convert to the more manageable nanomolar units:

$$[H^+ = 1.12 \times 10^{-7} \times 10^9 = \mathbf{112 \text{ nmol/L}}$$

N.B. 10^{-7} means "move the decimal point seven places to the left" whereas 10^9 means "move the decimal point nine places to the right". The net result is to move the decimal point 2 places (9-7 = 2) to the right so that 1.12 becomes 112.

b) If the hydrogen ion concentration is halved by the bicarbonate treatment then the new concentration will be 56 nmol/L or 0.56×10^{-7} mol/L (= 0.000000056 mol/L). Substitute this new value into the expression for pH:

$$pH = - \log_{10} [H^+] = - \log_{10} (0.000000056) = - (-7.25) = \mathbf{7.25}$$

Some calculators are cannot cope a large number of digits. Use can be made of the following property of logarithms:

$$\log (A \times B) = \log A + \log B$$

In *Q3(1)* to evaluate the log of 0.000000056 first write the number in exponential form

$$\log_{10} 0.000000056 = \log_{10} (5.6 \times 10^{-8})$$

evaluate each component separately then add to give the final result:

$$\log_{10} (5.6 \times 10^{-8}) = \log_{10} 5.6 + \log_{10} 10^{-8} = 0.75 + (-8) = - 7.25$$

Physiological hydrogen ion concentrations are always in the nanomolar range so a useful trick to simplify calculations is to keep the nonmolar term separate throughout the calculation.

Thus, in question *Q3(1)* above, an alternative expression for pH can be used:

$$\text{pH } 6.95 \quad = \quad 10^{-6.95}$$

Since 1 mol = 1,000,000,000 nmol, then 1 mol/L can be written as 10^9 mol/L. If concentration of hydrogen ions is expressed as nmol/L, then pH 6.95 can be written:

$$\text{pH } 6.95 \quad = \quad 10^{\,9-6.95} \quad = \quad 10^{\,2.05}$$

The antilog of 2.05 is 112 and so the hydrogen ion concentration is 112 nmol/L.

Similarly for part (b) where the concentration of hydrogen ions is halved to 56 nmol/L:

$$\text{pH} \quad = \quad -\log_{10}[H^+] \quad = \quad -\log_{10}(56 \times 10^{-9}) \quad = \quad -(\log_{10} 56 + \log_{10} 10^{-9})$$

Since $\log_{10}10^{-9}$ is - 9, then pH = $-(\log_{10}56 - 9) = -(1.75 - 9) = -(-7.25) = 7.25$

pH = - log $_{10}$ [H$^+$] ..... **Eq 3.3**

.

- Since the pH scale is logarithmic (to the base 10) a change in 1 pH unit corresponds to a 10-fold change in hydrogen ion concentration.
- Since the hydrogen ion concentration of pure water is 10^{-7} mol/L it follows that neutral pH is 7.
- pH changes in the opposite direction to hydrogen ion concentration. Solutions with pH greater than 7 are alkaline and solutions with pH less than 7 are acidic.
- The small 'p' means 'minus the logarithm to the base 10' and should not be confused with capital '*P*' which denotes partial pressure e.g. P_{CO_2}.

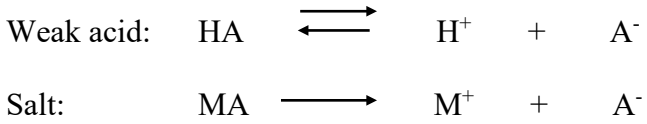
The expression for pH can be rearranged to calculate [H$^+$] from pH:

[H$^+$] = antilog $_{10}$ (- pH) ..... **Eq. 3.4**

Figure 3.3 Characteristics of pH

WHAT ARE BUFFERS? HOW DO THEY WORK?

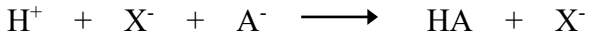
A buffer is a solution which resists change in pH when an acid or alkali is added. Many buffers consist of a mixture of a weak acid (HA) and its salt (MA). The salt can be regarded as completely ionised, but the weak acid is only partially ionised:



The dissociation constant of the weak acid (K_a) is given by:

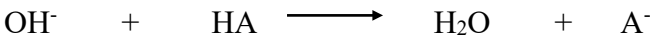
$$K_a = \frac{[H^+]\,[A^-]}{[HA]} \qquad \text{.....Eq.3.5}$$

The high concentration of the common anion (A^-) from the salt component suppresses ionisation of the weak acid (HA). The pH of the buffer solution will therefore depend on the relative amounts of weak acid and salt present. If a small amount of strong acid (HX) is added then the concentration of hydrogen ion will increase. To maintain constant K_a these hydrogen ions are "removed" by combination with the common anion (A^-) to form undissociated weak acid:



Hence the effect of the addition of the strong acid has been buffered. Consequently the strong acid anion (X^-) replaces the common anion (A^-) with only a small change in the $[A^-]/[HA]$ ratio so that pH does not change significantly.

If some alkali (XOH) is added then the hydroxyl ions (OH^-) react with, and therefore lower, the hydrogen ion concentration (with the formation of water). In order to maintain constant $[H^+][A^-]/[HA]$ ratio more weak acid dissociates so that the H^+ consumed by reaction with hydroxyl ion is replaced. This process can be considered as the titration of hydroxyl ions with weak acid (HA) so that the common anion (A^-) replaces the added hydroxyl ions:



Buffering results in a relatively constant pH but at the expense of altering the concentrations of the buffer components i.e. [HA] and $[A^-]$.

BUFFER CALCULATIONS

The expression for the dissociation constant of a weak acid (Eq.3.5) can be written slightly differently:

$$K_a = [H^+] \times \frac{[A^-]}{[HA]}$$

taking logarithms gives:

$$\log_{10} K_a = \log_{10} [H^+] + \log_{10} \frac{[A^-]}{[HA]}$$

Note that the logarithms of two numbers multiplied together is the same as the sum of their individual logarithms e.g. log (A x B) = log A + log B.

This equation can be rearranged by transferring $\log_{10}K_a$ to the right hand side (which then becomes negative) and $\log_{10} [H^+]$ to the left hand side (which also becomes negative):

$$-\log_{10} [H^+] = -\log_{10} K_a + \log_{10} \frac{[A^-]}{[HA]}$$

$-\log_{10} [H^+]$ is the same as pH and a term pK_a can be defined as $-\log_{10} K_a$, so that this equation becomes:

$$pH = pK_a + \log_{10} \frac{[salt]}{[\text{ acid}]} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{Eq3.6}$$

This is known as the *Henderson Hasselbalch* equation and can be used to calculate the amounts of salt and acid which must be used to prepare a buffer solution of a desired pH. Important quantitative properties of buffers are summarised in fig 3.4.

Question: Q3(2)

Calculate the amount in grams of lactic acid which must be added to 3 g of sodium hydroxide to give 1 litre of a solution with a pH of 4.5 (pK_a of lactic acid is 3.86; atomic weight of sodium is 23).

Answer to Q3(2)

The following reaction occurs between lactic acid (LactH) and sodium hydroxide:

$$LactH + NaOH \longrightarrow LactNa + H_2O$$

Both sodium lactate and lactic acid can dissociate to give lactate ions ($Lact^-$):

$$LactNa \longrightarrow Lact^- + Na^+$$

$$LactH \longrightarrow Lact^- + H^+$$

The relationship between the concentrations of lactate (a salt) and lactic acid (an acid) is governed by the Henderson Hasselbalch equation:

$$pH = pKa + \log_{10} \frac{[Lact^-]}{[LactH]}$$

In order to make use of this equation it is necessary to assume that the concentration of $Lact^-$ is equal to that of sodium, i.e.

1. That sodium lactate is completely dissociated, and

2. That the proportion of $Lact^-$ derived from lactic acid is negligible compared to that derived from sodium lactate. This proportion will be the same as the hydrogen ion concentration which can be calculated from the pH and is 0.0001 mol/L – clearly insignificant.

Next calculate the molar concentration of sodium hydroxide and use it in place of $[Lact^-]$:

Molecular weight (MW) of NaOH = 23 + 16 + 1 = 40

$$\text{Molar concentration of NaOH} = \frac{\text{Concentration (g/L)}}{\text{MW}} = \frac{3.0}{40} = 0.075 \text{ mol/L}$$

Reaction of sodium hydroxide with lactic acid (LactH) yields approximately an equal amount of lactate ($Lact^-$) ions:

$$NaOH + LactH \longrightarrow H_2O + Lact^- + Na^+$$

Substitute $Lact^-$ = 0.075, pH = 4.5 and pKa = 3.86 into the Henderson Hassebalch equation and solve for [LactH]:

$$4.5 = 3.86 + \log_{10} \frac{0.075}{[LactH]}$$

Rearranging gives:

$$4.5 - 3.86 = \log_{10} \frac{0.075}{[LactH]}$$

$$0.64 = \log_{10} \frac{0.075}{[LactH]}$$

taking antilogs of both sides:

$$\text{antilog } 0.64 = \frac{0.075}{[LactH]}$$

$$4.365 = \frac{0.075}{[LactH]}$$

Rearranging:

$$[LactH] = \frac{0.075}{4.365} = 0.0172 \text{ mol/L}$$

The total lactate needed is the sum of the ionised and unionised acid:

Total [Lact] = [LactH] + $[Lact^-]$ = 0.0172 + 0.075 = 0.0922 mol/L

Since all of this lactate must originate from the lactic acid added to the sodium hydroxide, calculate the weight of lactic acid required to give a concentration of 0.0922 mol/L:

Wt lactic acid (g) = Concentration (mol/L) x MW

MW of lactic acid (formula $C_3H_6O_3$) = (3 x 12) + (6 x 1) + (3 x 16) = 90

Wt of lactic acid required = 0.0922 x 90 = **8.30 g** (2 sig figs)

Question: Q3(3)

A buffer is required for a chromatographic procedure which has a pH of 7.0 and a total phosphate concentration of 0.1 mol/L. Calculate the amounts of anhydrous sodium dihydrogen phosphate and disodium hydrogen phosphate which need to be weighed to produce 1 litre of buffer. The pK_a of the dissociation is 6.82. (Atomic weights: Na = 23, P = 31).

Buffers are a mixture of a weak acid and its salt, the pH of which is defined by the *Henderson Hasselbalch* equation:

$$\mathbf{pH} = \mathbf{pK_a} + \mathbf{\log_{10}} \frac{\mathbf{[salt]}}{\mathbf{[acid]}} \quad \text{..... Eq.3.6}$$

- pH depends on the ratio $\frac{[salt]}{[acid]}$

- when [salt] = [acid], $\frac{[salt]}{[acid]} = 1$, $\log_{10} 1 = 0$ so pH = pK_a

  when $\frac{[salt]}{[acid]} = 0.1$, $\log_{10} 0.1 = -1$ so pH = $pK_a - 1$

  when $\frac{[salt]}{[acid]} = 10$, $\log_{10} 10 = 1$ so pH = $pK_a + 1$

Buffering capacity (i.e. the ability to resist a change in pH) is maximal when pH = pK_a. Outside the pH range pKa - 1 to pKa + 1 buffering action is minimal since the further the [salt]/[acid] ratio is from 1 the greater the resulting change in the logarithm of this ratio and hence pH when the same amount of acid or alkali is added.

- Outside 2 pH units either side of the pK_a the solution can be considered, for practical purposes, to consist entirely of acid or salt.

- For a weak base (B) the pK_a describes the dissociation of its *conjugate acid:*

$$BH^+ \rightleftharpoons B + H^+$$

Therefore the Henderson Hasselbalch equation can also be applied to a buffer consisting of a weak base and its salt since [salt] = [B] and [acid] = $[BH^+]$.

Figure 3.4 Quantitative properties of buffers

Answer to Q3(3)

The dissociation being considered is:

$$\underset{\text{dihydrogen phosphate ion}}{H_2PO_4^-} \rightleftharpoons \underset{\text{hydrogen phosphate ion}}{HPO_4^{2-}} + H^+$$

Phosphoric acid is a tribasic acid (with 3 pKa values) but at near neutral pH the contributions from the other species (H_3PO_4 and PO_4^{3-}) are negligible.

Substituting [salt] = $[HPO_4^{2-}]$, [acid] = $[H_2PO_4^-]$, $pK_a = 6.82$ and pH = 7.0 into the Henderson Hasselbalch equation gives:

$$7.0 = 6.82 + \log_{10} \frac{[HPO_4^{2-}]}{[H_2PO_4^-]}$$

Rearranging gives:

$$\log_{10} \frac{[HPO_4^{2-}]}{[H_2PO_4^-]} = 7.0 - 6.82 = 0.18$$

Taking antilogs:

$$\frac{[HPO_4^{2-}]}{[H_2PO_4^-]} = \text{antilog}_{10}\, 0.18 = 1.51 \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{Eq.3.7}$$

This equation now contains two unknowns and so at first sight appears impossible to solve. However, a further piece of information is given, namely, that the total phosphate concentration must be 0.1 mol/L. There are only two forms of phosphate to consider, therefore:

$$[HPO_4^{2-}] + [H_2PO_4^-] = [\text{total phosphate}] = 0.1 \text{ mol/L}$$

Rearranging to obtain an expression for one phosphate species in terms of the other (it doesn't matter which one) gives

$$[HPO_4^{2-}] = 0.1 - [H_2PO_4^-] \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{Eq.3.8}$$

which can be substituted into equation Eq. 3.7 to give an expression containing one variable which can then be solved:

$$\frac{0.1 - [H_2PO_4^-]}{[H_2PO_4^-]} = 1.51$$

$$0.1 - [H_2PO_4^-] = 1.51\,[H_2PO_4^-]$$

$$0.1 = 1.51\,[H_2PO_4^-] + [H_2PO_4^-] = 2.51\,[H_2PO_4^-]$$

$$[H_2PO_4^-] = \frac{0.1}{2.51} = 0.040 \text{ mol/L}$$

The other unknown, $[HPO_4^{2-}]$, can be obtained by substituting $[H_2PO_4^-] = 0.040$ into equation Eq.38:

$$0.1 = [HPO_4^{2-}] + 0.040$$

$$[HPO_4^{2-}] = 0.1 - 0.040 = 0.060 \text{ mol/L}$$

Next calculate the weight required for each phosphate salt:

$$\text{concn (g/L)} = \text{concn (mol/L)} \times \text{MW}$$

For anhydrous sodium dihydrogen phosphate (NaH_2PO_4):

$$\text{MW} = 23 + (2 \times 1) + 31 + (4 \times 16) = 120$$

$$\text{Wt required per litre} = 0.040 \times 120 = \mathbf{4.80\ g}$$

For anhydrous disodium hydrogen phosphate (Na_2HPO_4):

$$\text{MW} = (2 \times 23) + 1 + 31 + (4 \times 16) = 142$$

$$\text{Wt required per litre} = 0.060 \times 142 = \mathbf{8.52\ g}$$

THE BICARBONATE BUFFER SYSTEM

Carbonic acid (H_2CO_3) dissociates into hydrogen and bicarbonate ions:

$$H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$

Since carbonic acid is a weak acid, equilibrium is in favour of undissociated acid. Carbonic acid arises by the reaction of dissolved carbon dioxide with water, a reaction catalysed by the enzyme *carbonate dehydratase* (CD):

$$H_2O + CO_2 \rightleftharpoons H_2CO_3$$

These two reactions can be linked together (each with their own equilibrium constants, K_1 and K_2):

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$

The bicarbonate buffer system is central to acid-base homeostasis for two reasons:

- It affords a means of transporting the very insoluble CO_2 generated in tissues to the lungs where it can be eliminated. This is facilitated by the high concentration of carbonate dehydratase in erythrocytes which ensures rapid conversion of carbon dioxide to bicarbonate and *vice versa.* The hydrogen ions released when bicarbonate is generated are buffered by the high concentrations of haemoglobin also present in erythrocytes.

- It is an open system. The two components feeding into the system (carbon dioxide and bicarbonate ions) can be generated by the lungs and kidney respectively to ensure a constant CO_2/HCO_3^- ratio, and hence pH.

The equilibrium constants for these two reactions involving carbonic acid are given by:

$$K_1 = \frac{[H_2CO_3]}{[H_2O]\,[CO_2]} \quad \text{and} \quad K_2 = \frac{[H^+]\,[HCO_3^-]}{[H_2CO_3]}$$

In practice it is not possible to measure the very low concentrations of carbonic acid present in blood directly. Instead carbon dioxide is measured and so the above two equations are combined so that the carbonic acid term is eliminated and the buffer system can be described in terms of carbon dioxide and bicarbonate ion terms only:

Rearranging the first term gives an expression for carbonic acid concentration:

$$[H_2CO_3] = K_1[H_2O]\,[CO_2]$$

This expression can then be substituted for $[H_2CO_3]$ in the expression for K2:

$$K_2 = \frac{[H^+]\,[HCO_3^-]}{K_1\,[H_2O]\,[CO_2]}$$

Which can be rearranged to give an expression for $[H^+]$:

$$[H^+] = \frac{K_1K_2\,[H_2O\,\,[CO_2]}{[HCO_3^-]}$$

The concentration of water can be considered constant and combined with K_1 and K_2 to give a new constant, K_1'

$$[H^+] = \frac{K_1'\,[CO_2]}{[HCO_3^-]} \quad \text{..... Eq.3.9}$$

In routine practice the CO_2 content of blood is expressed as its partial pressure, Pco_2. *Pco_2 is the partial pressure of a gaseous phase which is in equilibrium with the sample.* This practice has arisen because calibrants are prepared by equilibrating blood with gaseous mixtures with known CO_2 content. *Henry's Law states that the amount of gas physically dissolved in a solution is proportional to the partial pressure of that gas.* The constant of proportionality is the *Bunsen solubility coefficient, α:*

$$[CO_2] = \alpha\ Pco_2$$

If Pco_2 is expressed in kiloPascals (kPa) this constant is 0.225, if it expressed in mm of Hg then it is 0.03. *One Pascal* (the SI unit of pressure) is the pressure exerted by 1 Newton acting on an area of 1 square metre. The Newton is the SI unit of force. One Newton is the force required to give a mass of 1 kg an acceleration of 1 metre per second per second.

By substituting αPco_2 for $[CO_2]$ in equation Eq.3.9 the following expression is obtained:

$$[H^+] = \frac{K_1'\ \alpha\ Pco_2}{[HCO_3^-]} \quad \text{.....Eq.3.10}$$

The constant K_1' can be combined with the solubility coefficient, α, to give a new constant. If $[H^+]$ is expressed in nmol/L, $[HCO_3^-]$ in mmol/L and Pco_2 in kPa then this constant is approximately 180:

$$[H^+] = \frac{180\ \ Pco_2}{[HCO_3^-]} \quad \text{..... Eq.3.11}$$

if Pco_2 is measured in mm Hg, then the value of the constant is 24.

Alternatively, if equation 3.10 is inverted and logarithms of both sides taken, then it becomes:

$$\log_{10}\frac{1}{[H^+]} = \log_{10}\frac{1}{K_1'} + \log_{10}\frac{[HCO_3^-]}{\alpha\ Pco_2}$$

By definition $\log_{10}\ 1/\ [H^+]$ is the pH and $\log_{10}\ 1/\ K_1'$ is the pK_1' and if these values are substituted into the above equation then the result is the familiar Henderson Hasselbalch equation. If 6.1 is substituted for pK_1' , and Pco_2 is expressed in kPa, then it becomes:

$$pH = 6.1 + \log_{10}\frac{[HCO_3^-]}{0.225\ Pco_2} \quad \text{..... Eq.3.12}$$

Both equations Eq.3.11 and Eq.3.12 contain three variables. Change in one variable must be accompanied by a change in at least one other. This is of importance for two reasons:

- Blood gas instruments only measure two variables (hydrogen ion concentration/pH and Pco_2). The third variable, the actual bicarbonate concentration [HCO_3^-], is calculated by substituting pH and Pco_2 in to the Henderson Hasselbalch equation.

- It is impossible to carry out experiments in which only one component is varied in order, for example, to investigate the susceptibility of chemoreceptors to single acid-base parameters.

It is important to remember that constants such as pK and α depend on other variables e.g. temperature.

In recent years there has been a move towards expressing the acid-base status of blood in terms of concentration using *nanomolar* units. The normal pH (7.4) then becomes 40 nmol/L (antilog$_{10}$ (-7.4) = 4.0 x 10^{-8} mol/L = 40 nmol/L) – quite an easy number to manage. The advantages of using hydrogen ion concentration instead of pH are:

- It allows a more intuitive approach i.e. acidosis is associated with an increase in [H^+] rather than a fall in pH

- The changes are linear. If pH is used then [H^+] has to increase 10-fold before the pH increases by a value of 1

- Unlike pH, [H^+] is linearly related to both Pco_2 and [HCO_3^-]. This makes it easier to calculate an expected compensatory change and helps interpretation of patients' blood gas results.

There is little doubt that using concentration is simpler than using pH (otherwise the three variables would have three very different types of units). Adoption of [CO_2] instead of Pco_2 would further simplify matters since all components would then be expressed as concentrations!

Question Q3(4)

The SHO in ITU carried out a blood gas analysis but failed to record all of the results in the patient's notes. The only available results are:

H^+ concentration	=	93 nmol/L
Actual bicarbonate	=	21 mmol/L

Calculate the pH, Pco_2 (in kPa) and carbon dioxide concentration (in mmol/L). Assume that the solubility coefficient of CO_2 (in kPa) is 0.225.

Answer Q3(4)

Since there are 1,000,000,000 (i.e. 10^9) nmol in a mol, then 93 nmol/L can be written as 93 x 10^{-9} mol/L.

Since pH = - $\log_{10}$ [H^+] and [H^+] = 93 x 10^{-9}, then

pH = -$\log_{10}$ (93 x 10^{-9}), which can be written pH = - [$\log_{10}$93 + $\log_{10}$ 10^{-9}]

Since $\log_{10}$93 = 1.97 and $\log_{10}$ 10^{-9} = -9

pH = - [1.97 + (-9)] = 9 – 1.97 = **7.03**

The Henderson-Hasselbalch equation for the HCO_3^-/CO_2 pair is:

$$pH = pKa + \log_{10} \frac{[HCO_3^-]}{\alpha\, Pco_2}$$

substitute: pH = 7.03, pKa = 6.1, [HCO_3^-] = 21 mmol/L , α = 0.225 then solve for Pco_2

$$7.03 = 6.1 + \log_{10} \frac{21}{0.225\, Pco_2}$$

$$7.03 - 6.1 = \log_{10} \frac{21}{0.225\, Pco_2}$$

$$0.93 = \log_{10} \frac{21}{0.225\, Pco_2}$$

$$\text{antilog}_{10}\, 0.93 = \frac{21}{0.225\, Pco_2}$$

$$Pco_2 = \frac{21}{0.225\ \text{antilog}_{10}\, 0.93} = \frac{21}{0.225 \times 8.51} = \frac{21}{1.92} = \mathbf{10.9\ kPa}$$

[CO_2] is calculated from the Pco_2 and α:

[CO_2] = α Pco_2 = 0.225 x 10.9 = **2.45 mmol/L**

URINARY BUFFERS

The maximum hydrogen ion gradient from tubular lumen to blood which can be generated by the kidney tubule is approximately 600:1. Since the hydrogen ion concentration of normal blood is about 40 nmol/L (pH = 7.4) this means that the maximum hydrogen ion concentration in urine will be 40 x 600 = 24000 nmol/L, which corresponds to a pH of 4.62.

The human body normally produces approximately 70 mmol of hydrogen ions per day of fixed acid (i.e. from sources other than carbon dioxide) which is excreted almost exclusively in the urine. The vast majority of these hydrogen ions are buffered by base in the urine, principally phosphate and ammonia:

$$H^+ + HPO_4^{2-} \rightleftharpoons H_2PO_4^- \qquad pK_a = 6.8$$

$$H^+ + NH_3 \rightleftharpoons NH_4^+ \qquad pK_a = 9.8$$

The high pK_a of ammonia means that at physiological pH and below the majority exists as ammonium ions. For this reason it has been debated whether or not excretion of urinary ammonium ions truly reflects excretion of acid. However, unlike phosphate, where its availability in the tubular fluid is relatively independent of hydrogen ion status, ammoniogenesis increases during acidosis and is therefore closely linked to acid-base status.

A quantity called the titratable acidity can be obtained by titrating urine with alkali back to the pH of blood (7.4). It therefore approximates to:

Titratable acidity (mol/L) = $[H^+]$ buffered by phosphate + free unbuffered $[H^+]$

Whereas the total acid excretion includes hydrogen ions buffered with ammonia:

[Total acid excretion] = [Titratable acidity] + $[NH_4^+]$

Question: Q3(5)

Over a 24 h period a patient excretes 1 litre of urine with a pH of 5.5, containing 50 mmol of inorganic phosphate and 10 mmol of ammonium. Assuming the blood pH is 7.40 and that the pK_{a2} for phosphate is 6.82, calculate:

a) The amount of free hydrogen ion excreted
b) The amount of hydrogen ion buffered by phosphate and ammonium ions.

Answer to Q3(5)

a) pH meters measure free hydrogen ions. Therefore convert the urinary pH to hydrogen ion concentration:

$$pH = -\log_{10}[H^+]$$

rearranging gives: $[H^+] = \text{antilog}(-pH)$

substituting pH = 5.5: $[H^+] = \text{antilog}(-5.5)$

$$= -(-3.16 \times 10^{-6})$$

$$= 3.16 \times 10^{-6} \text{ mol/L} \quad (\text{approx } 0.003 \text{ mmol/L})$$

Since the urine volume is 1 L, then in 24 h approx **0.003 mmol** of acid is excreted as free hydrogen ions (H^+).

b) At the initial pH of the glomerular filtrate (7.40) a significant proportion of phosphate exists in the "acid " form (i.e. $H_2PO_4^-$). As the pH of the tubular fluid falls to that of urine (pH = 5.5) the proportion of this "acid" form increases as hydrogen ions are buffered. Therefore the amount of hydrogen ions buffered will be the difference:

[secreted H^+ buffered by phosphate] = $[H_2PO_4^-]_{pH = 5.5} - [H_2PO_4^-]_{pH = 7.40}$

The concentration of $H_2PO_4^-$ both before and after buffering the secreted hydrogen ions can be calculated using the Henderson Hasselbalch equation since the total phosphate concentration (50 mmol/L) is known.

At pH 7.40:

$$7.4 = 6.82 + \log_{10}\frac{[HPO_4^{2-}]}{[H_2PO_4^-]}$$

$$7.4 - 6.82 = \log_{10}\frac{[HPO_4^{2-}]}{[H_2PO_4^-]}$$

Since [total phosphate] = $[HPO_4^{2-}] + [H_2PO_4^-] = 50$ mmol/L

$$[HPO_4^{2-}] = 50 - [H_2PO_4^-]$$

Substituting for $[HPO_4^{2-}]$ in the Henderson Hasselbalch equation:

$$7.4 - 6.82 = \log_{10}\frac{(50 - [H_2PO_4^-])}{[H_2PO_4^-]}$$

$$\text{antilog}_{10}(7.4 - 6.82) = \frac{50 - [H_2PO_4^-]}{[H_2PO_4^-]}$$

$$3.80 = \frac{50 - [H_2PO_4^-]}{[H_2PO_4^-]}$$

$$3.80\,[H_2PO_4^-] = 50 - [H_2PO_4^-]$$

$$3.80\,[H_2PO_4^-] + [H_2PO_4^-] = 50$$

$$4.80\,[H_2PO_4^-] = 50$$

$$[H_2PO_4^-] = \frac{50}{4.80} = \textbf{10.4 mmol/L}$$

At pH 5.5:

$$5.5 = 6.82 + \log_{10}\frac{[HPO_4^{2-}]}{[H_2PO_4^-]}$$

The above calculation procedure used at pH 7.40 is repeated to give:

$$[H_2PO_4^-] = \textbf{33.3 mmol/L}$$

By subtracting the concentration of $H_2PO_4^-$ at pH 7.40 from this value the concentration of secreted hydrogen ion buffered by phosphate can be calculated:

$$[H^+ \text{ buffered by phosphate}] = 33.3 - 10.4 = \textbf{22.9 mmol/L}$$

Since ammonia is almost entirely present as ammonium ions at pH 7.4, lowering the pH to 5.5 will not alter the amount of hydrogen ions buffered. Therefore the concentration excreted as NH^+ is 10 mmol/L. Since the 24 urine volume is 1 L:

Total buffered H^+ in urine =

$$H^+ \text{ buffered by phosphate} + H^+ \text{ buffered by ammonia}$$

$$= 22.9 + 10$$

$$= \textbf{32.9 mmol} \quad (33 \text{ to 2 sig figs})$$

By comparison the amount excreted as free hydrogen ions (0.003 mmol) is insignificant.

FURTHER QUESTIONS

1. What is the pH of 0.5 per cent (w/v) hydrochloric acid (assume complete dissociation, atomic weight Cl = 35.5)?

2. The reference range for blood pH is often quoted as 7.35-7.45. Express this range in terms of nannomoles of hydrogen ion per litre.

3. If the pH of urine is 6.0 and of blood 7.40, what is the gradient of hydrogen ion concentrations across the tubular cell walls?

4. Determine the secondary dissociation constant of phosphoric acid if blood of pH 7.00 contains 12.85 mg disodium hydrogen orthophosphate and 6.88 mg sodium dihydrogen orthophosphate per 100 mL of plasma.

5. What weight of anhydrous sodium carbonate and sodium bicarbonate would be required to prepare 500 mL of 0.2 M buffer pH 10.7 (pK_a HCO_3^- = 10.3)?

6. Isotonic sodium lactate, pH 7.4, is commonly administered intravenously to combat metabolic acidosis. How many ml of concentrated lactic acid (85% w/w, density 1.2) and how many grams of anhydrous sodium lactate would be used to prepare 2.5 L of this solution (pK_a lactic acid = 3.86)?

7. A 24 h urine collection has a pH of 5.5 and total phosphate content of 65 mmol. If the arterial pH is 7.40 and the pK_a for phosphate is 6.80, how many millimoles of hydrogen ion are excreted as titratable acidity using HPO_4^{2-} as buffer?

8. A buffer solution (pH 4.74) contains acetic acid (0.1 mol/L) and sodium acetate (0.1 mol/L) i.e. it is a 0.2M acetate buffer. Calculate the pH after addition of 4 mL of 0.025 M hydrochloric acid to 10 mL of the buffer.

Chapter 4

Spectrophotometry

Basic principles

A photometer is a device for measuring the amount of light transmitted through (or absorbed by) a solution. The wavelength of light absorbed will depend on the chemical structure of the compound present in the solution. A photometer therefore consists of a light source which generates a beam of light which passes through a cell (cuvette) containing the solution under analysis; the light which is transmitted through the solution falls on photodetector and generates an electric current proportional to the intensity of the light, which is then translated into a reading. Maximum sensitivity and specificity is achieved if the beam of light reaching the sample cell is parallel and of a constant wavelength (i.e. is monochromatic) and is of the wavelength which gives maximum absorption (minimum transmission) of the light. To achieve this, instruments isolate a portion of spectrum of white light generated from the light source (usually a bulb) by placing a monochromator in the light path before reaching the sample cell. In simple filter photometers a glass filter is used, whereas spectrophotometers use a prism or diffraction grating. Further details can be found in the standard textbooks of analytical chemistry.

Consider an incident light beam with intensity I_o passing through a square cell containing a solution of a compound which absorbs light at the wavelength being used. The intensity of the light reaching the detector, I, will be less than I_o. However it is the fraction of light absorbed (or transmitted) which is related to the concentration of the compound of interest. The fraction of incident light reaching the detector, I/I_o, is known as the *transmittance* (T). If expressed as a percentage then the term *percentage transmittance* ($\%T$) is used:

$$T = \frac{I}{I_o} \qquad \text{and} \qquad \%T = \frac{I \times 100}{I_o} \quad \text{..... Eq.4.1}$$

Therefore measurement of the transmitted intensity, I, by itself is useless, we also need to know the intensity of the incident light, I_o. A measure of I_o can easily be made if the sample cell is removed from the light path. In practice, instead of taking a reading of I_o, the instrument is set to a transmittance of 100% or an absorbance of zero (see later for

definition of these terms). Other factors affect the value of I_o when a reading of the sample is made, including a small amount of incident light reflected by the surface of the cell, absorbed by the material of the cell and by the solvent and/or components of the reagent. Therefore instead of setting the instrument to 100 %T or zero absorbance with the cell compartment empty (i.e. against air) it is customary to use a "blank" consisting of a cell containing either solvent alone or reagent without the analytical sample added. In a *single beam spectrophotometer* the blank is inserted into the cell compartment and the instrument blank value set, then the sample inserted etc. This has the disadvantage that error will be introduced if the instrument "drifts". This difficulty is overcome in *double beam spectrophotometers* in which the light source is split into two equal beams, one passing through the blank or reference and the other through the sample position, enabling the blank to be monitored continuously. Note that T and %T are ratios and so do not have units.

The absorption laws

Light is absorbed only when a photon collides with a molecule. It is not surprising therefore that the chance of a photon of light colliding with a molecule in solution, and hence the amount of light absorbed, will depend on the concentration of the compound in solution and the path length or thickness of the cell. This simple concept gives rise to the two absorption laws:

Bouger's Law or Lambert's Law: The fraction of light absorbed is proportional to the thickness of the absorber.

Beer's Law: The fraction of light absorbed by a compound in solution is proportional to its concentration.

Taking as an example a solution of a compound at a concentration of 1 g/L which absorbs half of the light passing through it, then:

$$T = \frac{I}{I_o} = 0.5 \qquad \text{and} \qquad \%T = \frac{I \times 100}{I_o} = 50\%$$

If the concentration is increased by 1 g/L to 2 g/L (i.e. doubled) then a half of the light which would have been transmitted by the solution containing 1 g/L, will be absorbed with a result that only a quarter of I_o will reach the detector Therefore of the incident light intensity I_0, one half is transmitted after passing through a cell containing 1 g/L of the compound and a quarter after passing through a solution containing 2 g/L. If the concentration is increased by a further 1 g/L to 3 g/L then only an eighth is transmitted

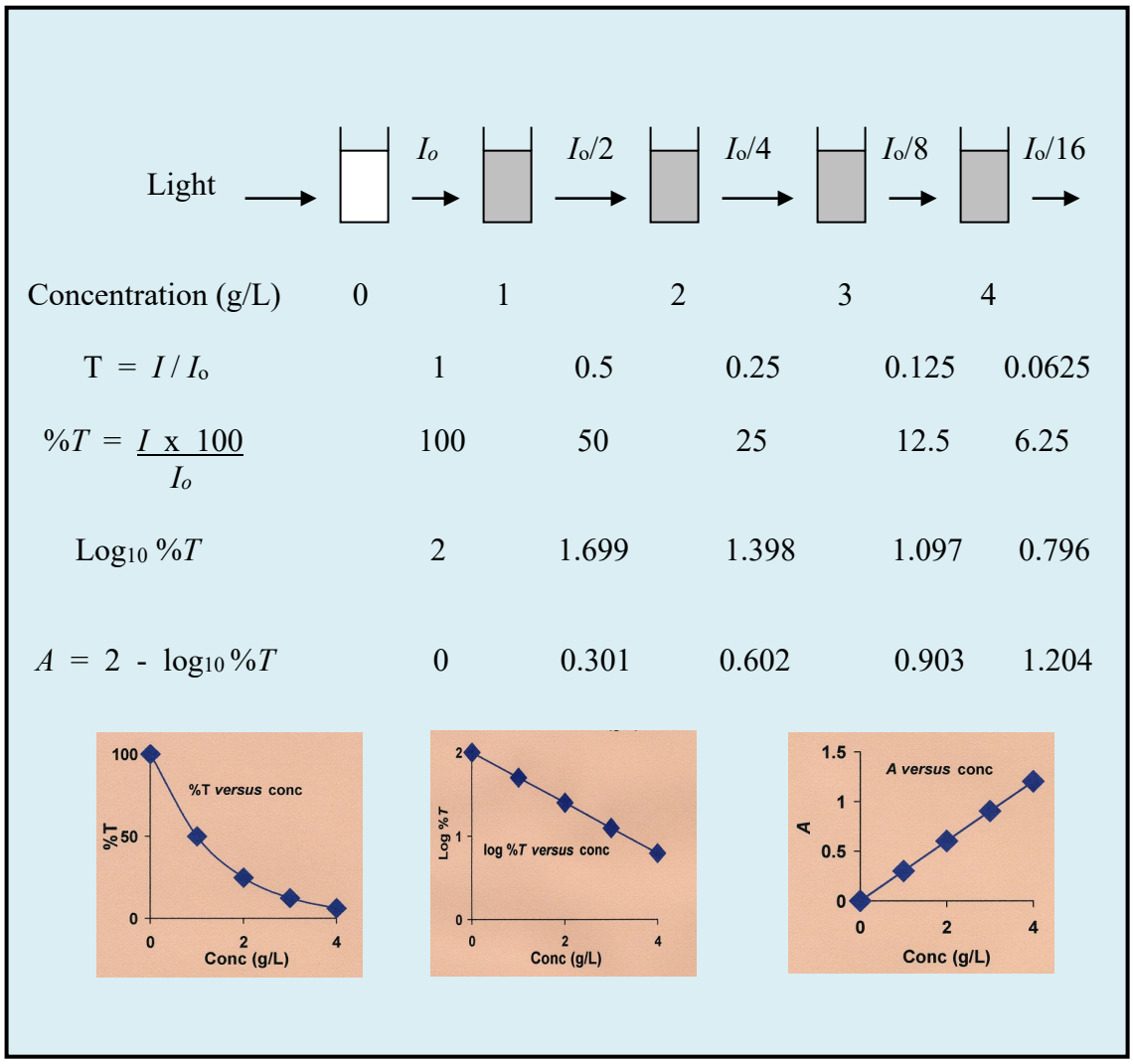


Figure 4.1 Relationship between % transmittance (%*T*), log_{10}% transmittance (log %*T*) and absorbance (*A*) illustrated by increasing the concentration of absorbing species in steps of 1 g/L

and if it is increased to 4 g/L, one sixteenth is transmitted. Therefore, of the incident light 1/2 is absorbed by 1 g/L, 3/4 by 2 g/L, 7/8 by 3 g/L and 15/16 by 4 g/L. It is clear that the relationship between concentration and transmittance is non-linear (Fig 4.1) and is in fact a *geometric* progression in which subsequent increase in concentration by 1 g/L decreases the transmitted light by a factor of 2. Taking logs of % transmittance (it doesn't matter to which base, but 10 is usually used) converts to a linear relationship with

concentration (logarithms are explained in chapter 3). However, at a concentration of zero, %T is 100%, $\log_{10}$%T is 2.0 and the value of $\log_{10}$%T decreases as the concentration of the absorbing species increases. It would be far more convenient if a value of zero was obtained when the concentration was zero, and the value then increased with increasing concentration. To achieve this an entity called *absorbance (A)* was created, which is the logarithm of the reciprocal of T:

$$A = \log_{10} \frac{1}{T} \quad \text{or} \quad A = \log_{10} \frac{I_o}{I} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{Eq.4.2}$$

Note that absorbance is a logarithmic function and so does not have units. Absorbance is some times called "*optical density*" and abbreviated "OD".

If however %T is used then %T can be converted to T by dividing by 100 to give %T/100 so that the above expression becomes:

$$A = \log_{10} \frac{100}{\%T}$$

Since $\log_{10} \frac{100}{T}$ can also be written as : $\log_{10} 100 - \log_{10} T$ and since $\log_{10} 100 = 2$ it follows that: $A = 2 - \log_{10} \%T$ …… Eq.4.3

This is a very convenient way to inter-convert A and %T. For example, when the concentration of the absorbing species is zero the %T is 100, the logarithm of 100 is 2 so that the absorbance must be 2 – 2 which is zero. When %T is 50, the logarithm of 50 is 1.699 so that the absorbance is 2 – 1.699 which is equal to 0.301. A plot of absorbance versus concentration is linear and passes through the origin (Fig 4.1). Similar reasoning shows that absorbance is also linearly related to cell path length. Therefore both Beer's and Lambert's Laws can be redefined as follows:

Beer's Law: Absorbance(A) is directly proportional to concentration(c).

Lambert's Law: Absorbance(A) is directly proportional to cell path length(b).

Question Q4(1)

A particular sample of a solution of a coloured substance, which is known to obey the Beer-Lambert Law shows 70% transmittance when measured in a 1 cm cell. Calculate the percent transmittance and absorbance of this solution if measured in a 0.5 cm cell and of a solution of twice the original concentration.

Answer Q4(1)

First calculate the absorbance of the original solution:

$$A = 2 - \log_{10} \%T = 2 - \log_{10} 70 = 2 - 1.845 = 0.155$$

According to Lambert's law, if the cell path length is halved by making the reading in a 0.5 cm cell, then the absorbance will be halved:

$$A \text{ in 0.5 cm cell} = \frac{0.155}{2} = \mathbf{0.078} \text{ (3 sig figs)}$$

Next calculate $\%T$ by substituting $A = 0.078$ into:

$$A = 2 - \log_{10} \%T$$

$$0.078 = 2 - \log_{10} \%T$$

Rearrange and evaluate:

$$\log_{10} \%T = 2 - 0.078 = 1.922$$

$$\%T = \text{antilog}_{10} 1.922 = \mathbf{84\%} \text{ (2 sig figs)}$$

If the initial concentration is doubled, then, according to Beer's law, the absorbance will also double:

$$A = 2 \times 0.155 = \mathbf{0.310}$$

Convert to %T as above:

$$0.310 = 2 - \log_{10} \%T$$

$$\log_{10} \%T = 2 - 0.310 = 1.69$$

$$\%T = \text{antilog}_{10} 1.69 = \mathbf{49} \% \text{ (3 sig figs)}$$

Combining these two laws gives: A is proportional to $b \times c$

And introducing a proportionality constant (a) gives the *Beer-Lambert equation:*

$$A = abc \quad \ldots\ldots\ldots \text{Eq.4.4}$$

Where *a* is a proportionality constant known as the *"absorptivity"*. The units for *a* will be the reciprocal of the units of *b* and *c*, and can be evaluated by substituting the units for A, b and c into equation Eq 4.4 then rearranging it. For example, if *b* is in cm and *c* is in mol/L (*A* of course, has no units), then the units of *a* will be $L.mol^{-1},cm^{-1}$:

$$A = a \times (mol/L) \times (cm)$$

$$a = \frac{A}{(mol/L) \times (cm)} = \frac{L}{mol \times cm} = L/mol/cm \ OR \ L.mol^{-1}.cm^{-1}$$

When concentration is in molar units, then *a* is termed the "*molar absorptivity,"* and the symbol ε is used. Terms and units used in spectrophotometry are defined in Fig 4.2.

NAME	SYMBOL	DEFINITION	UNITS
Transmittance	*T*	I/I_o	None
% Transmittance	%*T*	$I \times 100 / I_o$	%
Absorbance Optical density	*A* *OD*	$\log_{10} (I_o/I)$ "	None "
Molar absorptivity	ε	*A* when *c* = 1 mol/L, *b* =1cm	$L.mol^{-1}cm^{-1}$
Path length	*b*	Width of cell	cm
Wavelength	λ	Distance between light waves	nm
Absorption maximum	λ_{max}	λ at peak maximum	nm

Figure 4.2 Terms and units used in spectrophotometry. I_o = intensity of incident light (or reference), *I* = intensity of transmitted light, *c* = concentration

Question Q4(2)

The absorbance of a solution of pure bilirubin in chloroform when measured in a cuvette with a 0.5 cm path length (using a 0.5 cm cuvette containing chloroform to zero the instrument) is 0.268. If the concentration of bilirubin in the solution is 4 mg/L, calculate the molar absorptivity of chloroform. The molecular weight of bilirubin is 584.

Answer Q4(2)

$$A = a\ b\ c$$

Where: A = absorbance = 0.268

a = molar absorptivity = ?

b = cuvette path length = 0.5 cm

c = molar concentration =

$$= \frac{\text{concentration (mg/L)}}{1000 \times \text{Molecular weight}}$$

$$= \frac{4}{1000 \times 584} = 0.000\ 00685 \text{ mol/L} \quad (8.83 \times 10^{-6} \text{ mol/L})$$

Substituting these values:

$$0.268 = a \times 0.5 \times 0.000\ 00685 \quad \left(\text{units: cm} \times \frac{\text{mol}}{\text{L}}\right)$$

Rearranging and solving for a:

$$a = \frac{0.268}{0.5 \times 0.000\ 00685} \quad \left(\text{units: } \frac{\text{L}}{\text{mol}} \times \frac{1}{\text{cm}}\right)$$

$$= \mathbf{78{,}800\ L/mol/cm} \quad (\text{or } \text{L.mol}^{-1}\text{cm}^{-1})$$

Note that the answer is rounded to 3 significant figures because the absorbance measurement is only given to 3 decimal places.

How are absorbance measurements used to calculate the concentration of an analyte in a biological specimen?

Various approaches can be used, but all are based on the simple relationship between absorbance and concentration (Eq. 4.4):

$$A = abc$$

Since cuvettes with path lengths of 1 cm are almost universally used in spectrophotometric analysis, then the above expression can be simplified to:

$$A = ac$$

which can be rearranged to give:

$$c = \frac{A}{a} \quad \text{i.e. concentration} = \frac{\text{absorbance}}{\text{absorptivity}} \quad \ldots\ldots \text{Eq.4.5}$$

so that concentration can be calculated from an absorbance reading if the absorptivity is known. Note that the units of absorptivity should be appropriate for the units of concentration being used. For example, if concentration is in µmol/L then the micromolar absorptivity should be used (units: $L.\mu mol^{-1}cm^{-1}$). The assumption is also made that at zero concentration the absorbance is zero (i.e. the instrument is either zeroed on, or the measurement made with reference to, a cuvette of the same path length containing either the same solvent or the reagent being used.

1. How to use a documented value for absorptivity (*a*)

If the absorptivity of the absorbing species (that is, the analyte or a chromogen formed by the reaction of the analyte with a reagent) is known, then the concentration of the unknown can be calculated from its absorbance reading. It is vital to allow for any dilution involved and differences in concentration units.

Question 4(3):

0.1 mL of serum is mixed with 3.0 mL of a reagent which forms a coloured product with glucose. After the reaction has reached equilibrium the absorbance (versus a reagent blank) in a 1 cm cuvette was found to be 0.250. If the absorpivity of the chromogen is 933 $L.mol^{-1}cm^{-1}$ what is the serum glucose concentration expressed as mmol/L?

Answer Q4(3):

The expression relating absorbance to absorptivity (Eq. 4.5) is used:

$$c = \frac{A}{a}$$

Where c = glucose concentration in mol/L = ?
A = absorbance reading = 0.250
a = absorptivity of the chromogen = 933 $L.mol^{-1}cm^{-1}$

Substituting these values:

$$c = \frac{0.250}{933} \text{ mol/L}$$

Note that as the units of absorptivity of the chromogen are $L.mol^{-1}cm^{-1}$, the calculated concentration of glucose is in mol/L NOT mmol/L. If the above result is multiplied by 1000 (since there are 1000 mmol in a mol) then the result will be in mmol/L. Since the cuvette path length is also 1 cm then no correction need be made for differences in path length.

$$c = \frac{0.250}{933} \times 1000 \text{ mmol/L}$$

The absorptivity relates to the concentration of glucose derived chromogen in the solution of which the absorbance is being measured. However, 0.1 mL of serum was mixed with 3.0 mL reagent (i.e. diluted to 3.1 mL) before the absorbance reading was made. Therefore to convert the calculated concentration of chromogen (c) to glucose concentration in the undiluted sample, it is multiplied by 3.1 and divided by 0.1:

$$\text{Serum glucose} = \frac{0.250}{933} \times 1000 \times \frac{3.1}{0.1} = \mathbf{8.3 \text{ mmol/L}}$$

Note that the absorbance reading is multiplied by the reciprocal of absorptivity. Therefore if a large number of analyses are to be carried out then it is relatively easy to program this factor into the memory of a pocket calculator so that concentrations may be read off directly upon entering absorbance readings. Alternatively, using the spreadsheet facility of a PC, a table can be generated with absorbance values and their corresponding concentrations. It is important to allow for any differences in units and any dilution of the biological sample. If *question 4(3)* is taken as an example:

$$c = A \times \frac{1}{933} \times 1000 \times \frac{3.1}{0.1} \text{ mmol/L}$$

$$c = A \times 33.2 \text{ mmol/L}$$

so that multiplication of the absorbance reading by 33.2 gives the serum glucose concentration in mmol/L.

2. Using a standard solution

Suppose we have two samples: sample 1 with concentration c_1 gives rise to an absorbance of A_1, sample 2 with a concentration c_2 gives rise to absorbance A_2. Then we can write two relationships for absorptivity:

$$a = \frac{A_1}{c_1} \quad \text{and} \quad a = \frac{A_2}{c_2}$$

Since absorptivity (a) is constant then $\frac{A_1}{c_1} = \frac{A_2}{c_2}$

which can be rearranged to: $c_2 = \frac{A_2}{A_1} \times c_1$

If solution 1 is a standard solution in which the concentration of the analyte being measured is accurately known, and solution 2 is the biological specimen in which the concentration of the analyte is unknown, then:

$$\text{Concentration}_{\text{unknown}} = \frac{\text{Absorbance}_{\text{unknown}} \times \text{Concentration}_{\text{standard}}}{\text{Absorbance}_{\text{standard}}} \quad \ldots\ldots\ldots \text{Eq.4.6}$$

It is important that the absorbances are measured against an appropriate blank.

Question Q4(4):

A method for the measurement of serum glucose involves adding 0.1 mL of sample (serum, water or standard) to 3 mL of reagent, then after 10 min incubation at room temperature, measuring the absorbance at 500 nm in a cuvette with a 1 cm path length using an identical cuvette containing distilled water as reference. The readings using serum, standard or water as sample were 0.302, 0.353 and 0.052 respectively. If the concentration of glucose in the standard is 10 mmol/L, calculate the glucose concentration in the serum.

Answer Q4(4):

Since all the absorbances are measured against a cuvette containing water as reference, the first step is to subtract the absorbance of the blank (in which 0.1 mL of distilled water is used as the sample) so that in the absence of analyte the absorbance reading will be zero:

Sample	*A* (versus water)	*A* (corrected for blank)
Blank	0.052	0.000
Standard	0.353	0.301
Serum	0.302	0.250

Since the glucose concentration in the standard is known to be 10 mmol/L

$$\text{Serum glucose (mmol/L)} = \frac{A_{\text{serum}}\ \text{(corrected)} \times 10}{A_{\text{standard}}\ \text{(corrected)}}$$

$$= \frac{0.250 \times 10}{0.301} = \mathbf{8.3\ mmol/L}$$

Failure to correct each reading for the absorbance reading of the blank (0.052) would result in an erroneous answer of 8.6 mmol/L (0.302 x 10 / 0.353).

Since the dilutions of the sample (serum or standard) with reagent are identical, the dilutions cancel and do not need to be included in the calculation. However, any dilution of the sample which is *different* to that of the standard *must* be taken into account.

3. Using a calibration curve

If Beer's Law is obeyed and cuvettes with a constant path length are used then the equation relating absorbance to concentration (Eq. 4.4) is a linear function:

$$A = ac$$

In other words if, for a series of solutions, absorbance (A) is plotted on the vertical axis and concentration (c) on the horizontal axis then a straight line is obtained which passes

through the origin and has a slope (i.e. gradient) of a (Fig 4.3). Such a plot is called a *calibration curve*. There are various ways in which a calibration curve can be used to calculate the concentrations of analyte in an unknown sample:

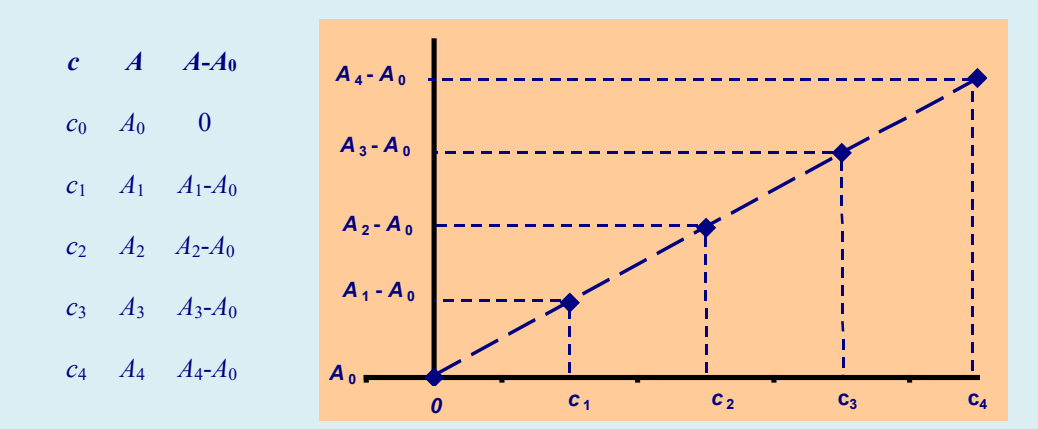


Figure 4.3 Construction of a calibration curve from a set of absorbance readings obtained for a series of standard solutions. c_0 is a reagent blank (standard with zero concentration) with an absorbance of A_0

Question Q4(5)

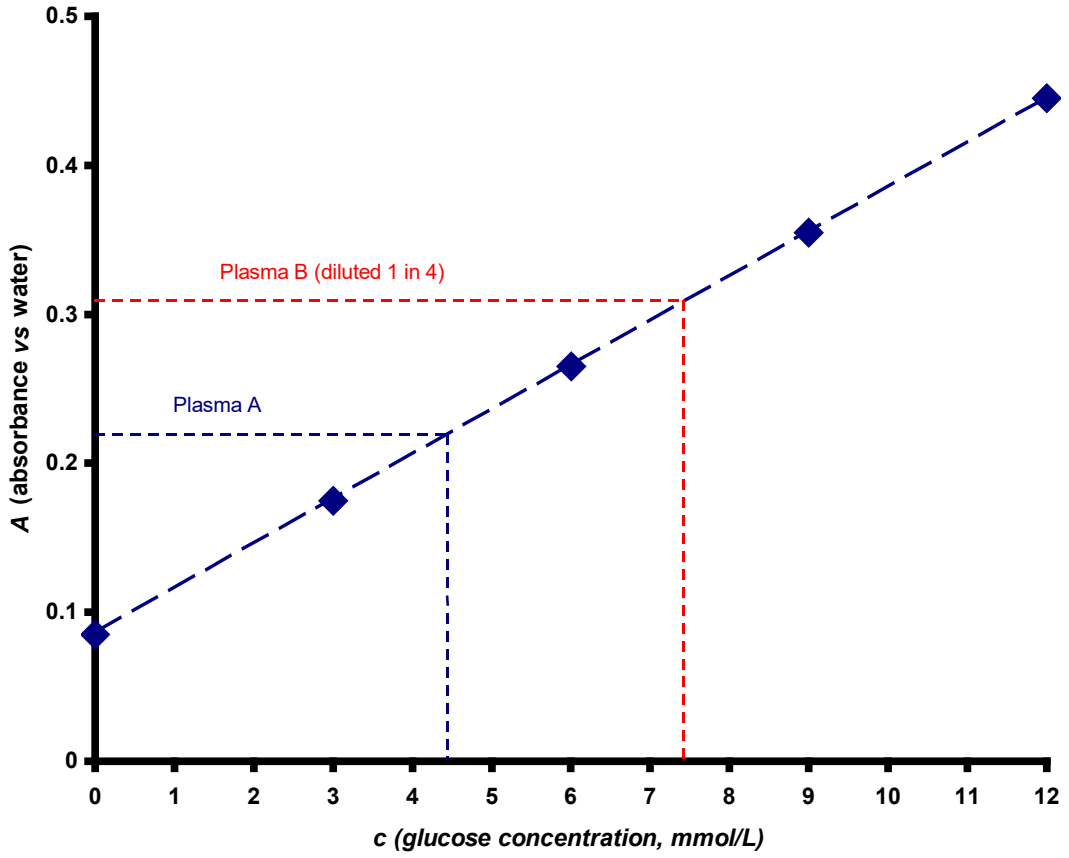
A manual method for the determination of plasma glucose involves mixing 0.1 mL of sample (water, glucose standard solution or plasma sample for the reagent blank, standard and unknown sample respectively) with 3 mL of reagent, then after 10 min incubation at room temperature, measuring the absorbance of the chromogen at 500 nm. The following results were obtained:

Sample	Absorbance
Water (zero standard)	0.085
Standard, 3 mmol/L	0.175
" , 6 "	0.265
" , 9 "	0.355
" , 12 "	0.445
Plasma A	0.220
" B	0.800
" " (1 in 4 dilution)	0.246

Before the absorbance measurements were made, the instrument was set to zero using an identical 1 cm cuvette containing distilled water. Determine the concentrations of glucose in plasma samples A and B.

Answer Q4(5):

There are several approaches which can be used. The simplest is to plot the data with absorbance (*A*) on the vertical axis glucose concentration (*c*) on the horizontal axis:

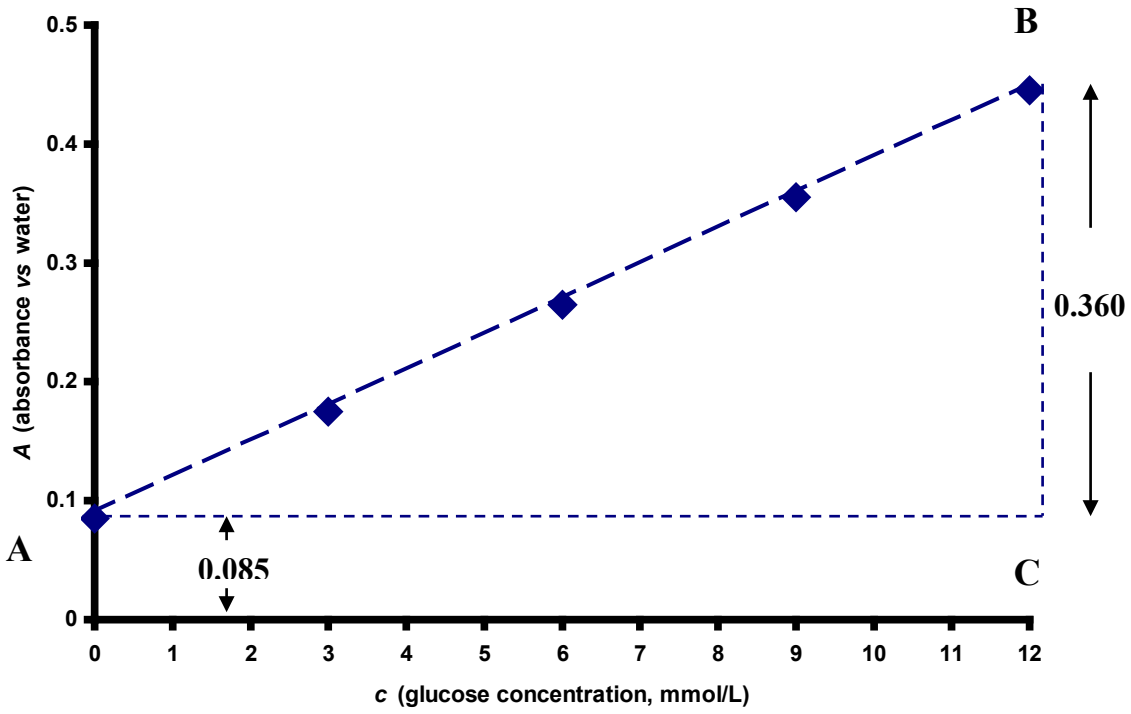


A straight line is then drawn through the points in such a way as to produce the best visual fit to the data. It is then an easy matter to read off the values for the plasma samples from this curve.

For plasma A (absorbance = 0.220) a horizontal line is drawn from the absorbance value on the vertical scale until it intersects the calibration curve. From this point a vertical line is drawn downwards until it intersects the horizontal axis. The reading at this point gives the glucose concentration in the sample as **4.5 mmol/L.**

This procedure cannot be followed for the neat (undiluted) absorbance value for plasma B (0.80) since this value is beyond the range of the calibration curve i.e. the absorbance value of the highest standard is only 0.445. Using a ruler the calibration curve could be extrapolated to this value but this would assume that the calibration is linear to this point. This is unlikely to be true if either the linear range of the instrument is exceeded or the capacity of the reagent is exhausted. The absorbance for plasma B which had been previously diluted 1 in 4 is well within the working range of the assay (0.310) and reading the plasma glucose concentration from the calibration curve using an identical procedure as for plasma A gives a value of 7.5 mmol/L. Since the 0.1 mL of plasma B sample had been diluted 1 in 4 prior to assay, this result is multiplied by 4 to give a final glucose result of 30 mmol/L.

Another approach is to determine the equation which describes the line through the standards (the calibration curve) then use it to arithmetically calculate the concentration of glucose in the plasma sample. The first step is to draw a right angled triangle (ABC) for which the calibration curve is the hypotenuse:



The linear equation describing the calibration curve takes the form:

$$A = \text{Intercept} + (c \times \text{slope})$$

The "intercept" is the point at which the calibration curve crosses the vertical axis i.e. the value of the absorbance (A) when glucose concentration (c) equals zero, and is 0.085. The slope is the change in absorbance which is observed for a 1mmol/L increment in glucose concentration i.e. BC divided by AC:

$$\text{Slope} = \frac{(0.445 - 0.085)}{12} = \frac{0.360}{12} = 0.030$$

So that the equation for the calibration curve is:

$$A = 0.085 + 0.03\,c$$

Which can be re-arranged to:

$$c = \frac{(A - 0.085)}{0.03}$$

By substituting absorbances obtained for the plasma samples, their glucose concentration is easily calculated:

$$\text{Plasma A glucose} = \frac{(0.220 - 0.085)}{0.03} = \mathbf{4.5\ mmol/L}$$

For plasma B the absorbance obtained from the diluted sample should again be used:

$$\text{Plasma B glucose (1 in 4 dilution)} = \frac{(0.310 - 0.085)}{0.03} = 7.5\ \text{mmol/L}$$

$$\text{Plasma B glucose (neat)} = 4 \times 7.5 = \mathbf{30\ mmol/L}$$

Some modern spectrophotometers can be calibrated directly as the blank and standard absorbances are read so that as unknown samples are read the result is displayed directly in concentration units rather than absorbance values. Again it is important to emphasize that this equation should only be used if the calibration curve is linear and for the concentration range covered by the standards. The glucose concentration can be calculated for the undiluted plasma B sample (A = 0.80) from the above equation:

$$\text{Plasma B glucose} = \frac{(0.800 - 0.085)}{0.03} = 23.8\ \text{mmol/L}$$

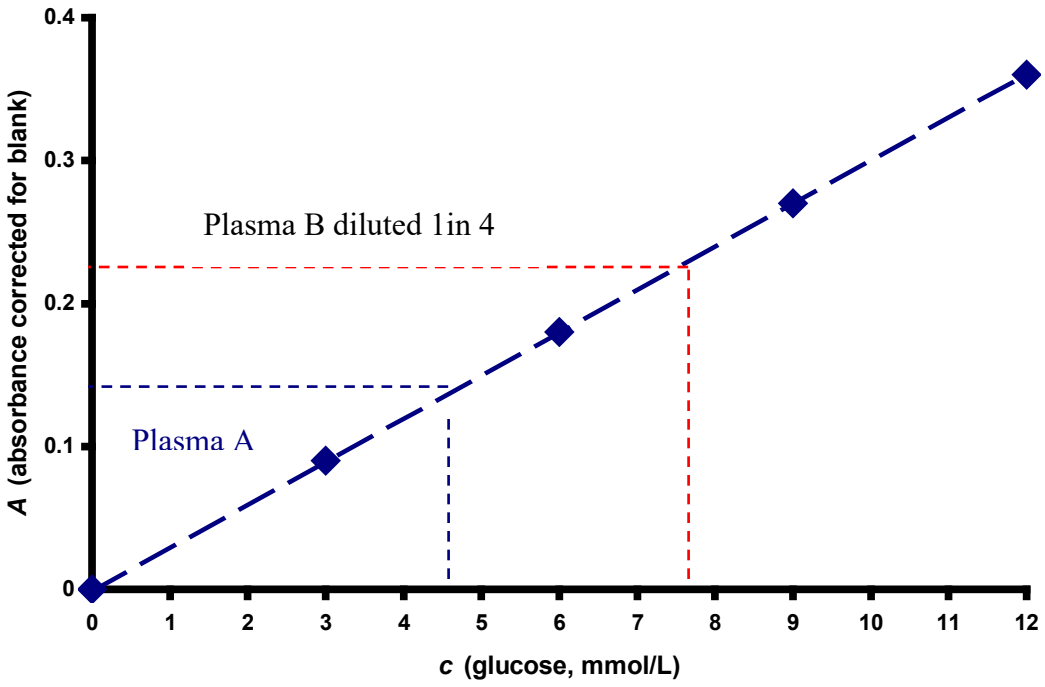
Clearly this value is erroneously low and illustrates the danger of extrapolating the standard curve beyond the highest standard used in its construction.

Sometimes the scatter of the points of the calibration curve make it difficult to decide where to draw the line of best fit. Under these circumstances the statistical line of best fit can be calculated using statistical methods. Modern instruments often use very complex mathematical techniques to calculate the best equation which describes the calibration curve.

An alternative approach is to subtract the reagent blank absorbance (the absorbance obtained when the glucose concentration of the standard is zero) form all readings (or alternatively make all readings using the reagent blank as reference): :

Sample	**Absorbance**	**Absorbance –Blank absorbance**
Water (zero standard)	0.085	0.000
Standard, 3 mmol/L	0.175	0.090
" , 6 "	0.265	0.180
" , 9 "	0.355	0.270
" , 12 "	0.445	0.360
Plasma A	0.220	0.135
" B	0.800	0.715
" " (1 in 4 dilution)	0.246	0.161

This has the effect of forcing the calibration curve through the origin:



Note that the slope is the same (0.03) although the intercept on the absorbance axis (previously 0.085) is now zero. Thus the equation describing the calibration curve is now simpler:

$$A \text{ (blank corrected)} = 0.03\,c$$

Which can easily be rearranged to

$$c = \frac{A \text{ (blank corrected)}}{0.03} \text{ mmol/L}$$

and used to calculate the glucose concentration for plasma samples. It is vital that the reagent blank is also subtracted from the absorbance readings for the plasma samples before their concentrations are either calculated from the above formula or read directly from the calibration curve.

An alternative approach when the absorbance of a plasma sample is beyond the working range of the calibration curve is to dilute the final reaction mixture, either with distilled water or reagent until the absorbance falls within range. This practice overcomes non-linearity due to the spectrophotometer but not if non-linearity is due to exhaustion of the reagent or one of its components. If the reaction mixture was diluted with more reagent then the correction to be applied is the same as if the assay had been repeated with diluted plasma sample. However, if the reaction mixture has been diluted with water then the reagent blank is no longer appropriate since the contribution of the reagent to the final absorbance reading has been reduced. The reagent blank to be subtracted should first be divided by the dilution factor used.

If the relationship between absorbance and concentration is linear (i.e. Beer's Law is obeyed) then the minimum number of data points required to determine the linear equation relating the two quantities is 2. The term *single point calibration* is often used which is, strictly speaking, incorrect. A second point is always required to define a straight line (the shortest distance between two points in the same plane). It is always inferred that this second point is the value when concentration is zero i.e. the blank. There are other approaches to using a set of standard absorbances. In general if there are n standards with their corresponding absorbances then the ratios of their absorbances to their concentrations are equal:

$$\frac{A_1}{c_1} = \frac{A_2}{c_2} = \frac{A_3}{c_3} = \frac{A_4}{c_4} = \ldots\ldots \frac{A_n}{c_n}$$

The only time this relationship does not hold is when the concentration (c_0) is zero, since any number divided by zero becomes infinite. A further requirement is that all measurements should be made using an appropriate blank as reference.

The individual ratios of A/c can be averaged to produce a mean value for A/c which can then be used to calculate the concentrations in unknown samples:

$$\text{Concentration of unknown} = \text{mean } (c/A)_{\text{standards}} \times A_{\text{unknown}}$$

Although this procedure is simple and takes account of the imprecision of absorbance measurements of the standards it is not recommended. The absorbances of the individual standards provide a further piece of information: confirmation that the relationship between absorbance and concentration is in fact linear. Therefore, a calibration curve should always be plotted or accessed mathematically.

Dealing with mixtures

Often we encounter solutions which contain more than one light-absorbing species. If their absorption spectra overlap then the presence of one of these species may interfere with the determination of the other. Fig 4.4 shows the absorption spectra of two compounds, A and B. At the wavelength of maximum absorption (λ_{max}) of either compound there is significant absorption of light by the other. The spectrum of a mixture of A and B at the same concentrations present in their separate solutions has an absorption maximum somewhere in between that of the individual compounds. Provided there is no interaction between the two compounds, their individual spectra are *additive*.

Beer's law states that absorbance (A) is equal to absorptivity (a) multiplied by concentration (c):

$$A = a\,c$$

Similar equations can be written for the two species A and B:

$$A_A = a_A \times c_A$$

$$A_B = a_B \times c_B$$

Provided there is no interaction between A and B, their absorbances are additive, so that the total absorbance (A_{total}) is the sum of their individual absorbances:

$$A_{total} = A_A + A_B$$

Which can be written in terms of their individual concentrations and absorptivities:

$$A_{total} = (a_A \times c_A) + (a_B \times c_B) \quad \ldots\ldots\ldots\ldots \text{Eq.4.7}$$

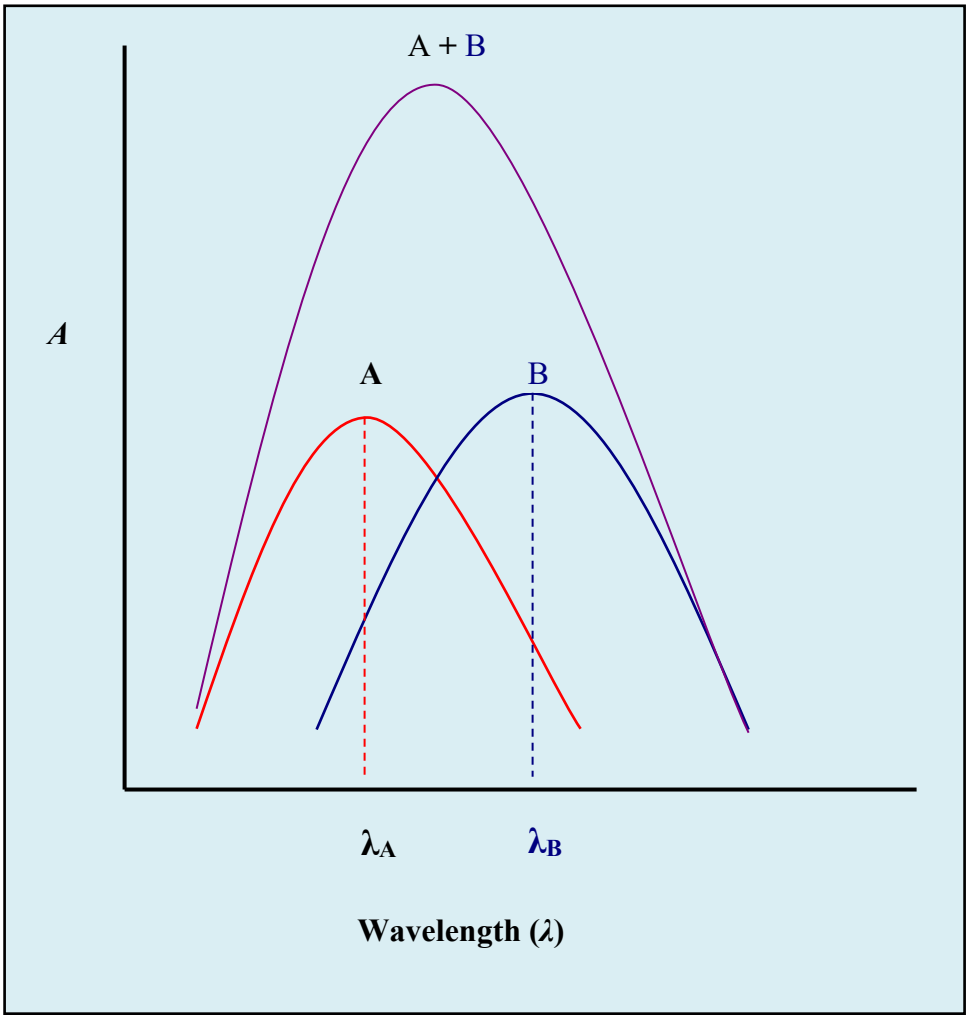


Figure 4.4 Over-lapping spectra of two compounds (A and B) and a mixture of both (A + B)

At first sight it may appear impossible to use the measured absorbances to calculate the individual concentrations of A and B in a mixture of both, even if the individual absorptivites of A and B are known i.e. the equation contains two unknowns. However, if the absorbance is also measured at a second wavelength (usually the λ_{max} of the other species) then another equation for the measured absorbance can be set up similar to Eq. 4.7.

In general, if there are two species, A and B

A has absorptivities of $a_{A\lambda1}$ at wavelength 1 and $a_{A\lambda2}$ at wavelength 2
B " " " $a_{B\lambda1}$ " " " " $a_{B\lambda2}$ " " "

A_1 is the measured absorption of the mixture at wavelength 1
A_2 " " " " " " " " " " 2

c_A is the concentration of A in the mixture
c_B " " " " B " " "

Then two equations can be written for the measured absorption, one for each wavelength:

$$A_1 = a_{A\lambda1}\, c_A + a_{B\lambda1}\, c_B$$
$$A_2 = a_{A\lambda2}\, c_A + a_{B\lambda2}\, c_B \qquad \text{..... Eq.4.8}$$

These form a set of 2 simultaneous equations (each containing the same two unknowns) which can then be solved in the usual way (see Fig 4.5).

The same principle can be applied to mixtures containing 3, 4 or even more components. Absorbance measurements must be made at the *same* number of wavelengths as the components in the mixture. However, the simultaneous equations become increasing complex and difficult to solve.

Question Q4(6)

A chromatographic method for the separation of 4 different drugs fails to completely resolve two of them (drugs A and B). Fortunately drugs A and B have overlapping spectra with the following absorptivities ($L.mol^{-1}cm^{-1}$):

	260 nm	**280 nm**
Drug A	100	500
Drug B	1000	200

Fractions from peaks A and B were pooled and the absorbance of the mixture measured in a cuvette with a 1 cm light path using solvent as reference. The absobance reading at 260 nm was 0.4 and at 280 nm 0.8. Calculate the individual concentrations of A and B in the pooled fractions.

Simultaneous equations are a set of two or more equations in two or more unknowns that are simultaneously true. For example, consider two equations in which x and y are unknown but in which the values of x and y are identical:

$$2x + 3y = 16 \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(i)}$$

$$3x + 2y = 14 \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(ii)}$$

Either or both equations are multiplied by factors in order to render *one* of the terms equal. Subtraction of one equation from the other then eliminates one of the variables. The resulting equation is then solved for the other variable. Simple inspection often suggests a suitable factor to use but an approach which always works is to multiply the first equation by the constant in front of one of the variables in the second equation and the second equation by the constant in front of the same variable in the first equation. In the above example if the *whole of* equation (i) is multiplied by 3 (the constant in front of x in equation (ii)) which then becomes equation (iii), and the *whole of* equation (ii) is multiplied by 2 (the constant in front of x in equation (i)), which then becomes equation (iv) then both equations contain $6x$. Subtraction of equation (iv) from (iii) eliminates variable x, and the result (equation (v)) can be solved for the other variable (y):

$$6x + 9y = 48 \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(iii)}$$

$$6x + 4y = 28 \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(iv)}$$

$$5y = 20 \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(v)}$$

$$y = \frac{20}{5} = \mathbf{4}$$

This value for y can then be substituted into either equation (i) or (ii) which is then solved for the other variable (x). Using equation (i):

$$2x + (3 \times 4) = 16$$

$$2x = 16 - (3 \times 4x) = 16 - 12 = 4$$

$$x = 4/2 = \mathbf{2}$$

Figure 4.5 A method for the solution of a pair of simultaneous equations containing two unknowns

Answer Q4(6):

At each wavelength the measured absorbance is equal to the sum of the absorbances due to dugs A and B. The absobance due to either drug is given by its molar concentration multiplied by its molar absorptivity at that particular wavelength. Therefore an eauation for the measured absorbance at each wavelength can be set up:

At 260 nm: $A_{260} = a_{A260}\, c_A + a_{B260}\, c_B$ ……..(i)

At 280 nm: $A_{280} = a_{A280}\, c_A + a_{B280}.c_B$ ……(ii)

Substituting values for the absorbances, molar concentrations and molar absorptivities into equations (i) and (ii):

At 260 nm: $0.4 = 100\, c_A + 1000\, c_B$ ……(i)

At 280 nm: $0.8 = 500\, c_A + 200\, c_B$ ……(ii)

Multiply equation (i) by 5 throughout (to become equation (iii)), then subtract equation (ii) from (iii) so as to eliminate the c_A term:

$2.0 = 500\, c_A + 5000\, c_B$ ……...(iii)

$0.8 = 500\, c_A + 200\, c_B$ ……(ii)

$1.2 = 4800\, c_B$

Solve for c_B:

$$c_B = \frac{1.2}{4800} = 0.00025 \text{ mol/L} = \mathbf{0.25 \text{ mmol/L}}$$

Substitute 0.00025 for c_B in equation (i), then solve for c_A:

$0.4 = 100\, c_A + (1000 \text{ x } 0.00025)$

$0.4 = 100\, c_A + 0.25$

$100\, c_A = 0.4 - 0.25 = 0.15$

$$c_A = \frac{0.15}{100 \text{ x}} = 0.0015 \text{ mol/L} = \mathbf{1.5 \text{ mmol/L}}$$

FURTHER QUESTIONS

1. An aqueous solution in a 1 cm cell has an absorbance of 0.23 when read against a water blank at 500 nm. Assuming Beer's Law is obeyed, what volume of this solution would need to be added to 100 mL of water to give a solution which absorbs 30% of the light entering it under the same measurement conditions?

2. Calculate the absorbances corresponding to the following percentage transmittance readings:

   a) 95 b) 75 c) 50 d) 25 e) 10 f) 1

3. Calculate the % of incident light transmitted by solutions with the following absorbances:

   a) 0.1 b) 0.25 c) 0.50 d) 0.75 e) 1.00 f) 2.00

4. A solution of a compound (concentration 100 mmol/L) was placed in a cuvette with a 1 cm light path and the percentage of incident light transmitted was 18.4. Calculate the molar absorptivity of the compound.

5. The transmittance of a solution of NADH at 340 nm is 45%. What is the absorbance at 340 nm of a 1 in 5 dilution of this solution?

6. 75 mg of faeces were homogenised in 1 mL of concentrated hydrochloric acid, then 3 mL diethylether added, mixed, 3 mL of water added and mixed again. After centrifugation the aqueous phase (volume 4.5 mL) was scanned in a spectrophotometer using a cell with a 1 cm pathlength and the peak height at 405 nm due to porphyrin, after applying a background correction, was 0.35 absorbance units. A separate 0.250 g portion of faeces was dried in a 100^{o}C oven for 3 hours after which it's weight was 0.125 g. Given that the molar absorption coefficient of porphyrin is 2.75 x 10^5 L/mol/cm calculate the porphyrin concentration in nmol/g dry weight of faeces.

7. A solution containing a substance of molecular weight 400 at a concentration of 3 g/L transmitted 75% of incident light of a particular wavelength in a 1 cm cuvette. Calculate the % of incident light of the same wavelength that would be transmitted by a solution of the same substance at a concentration of 4 g/L and calculate the molar absorption coefficient for that substance at this wavelength.

8. The absorbances of a solution containing NAD and NADH in a 1cm light path cuvette were 0.337 at 340 nm and 1.23 at 260 nm. The molar extinction coefficients are:

   NAD: 1.8×10^4 at 260 nm, 1.0×10^{-3} at 340 nm
   NADH: 1.5×10^4 at 260 nm, 6.3×10^3 at 340 nm

   Calculate the concentrations of NAD and NADH in the solution.

9. 25 mg of bilirubin ($C_{33}H_{36}O_6N_4$) were dissolved in 4 mL of dimethyl sulphoxide; 200 µL of this solution was diluted to 250 mL with chloroform. This solution gave an absorbance of 0.502 when measured in a 1 cm cell against a chloroform blank.

   Given that the molar absorptivity of bilirubin under these conditions is 6.07×10^4, calculate the percentage purity of the bilirubin.

10. A method for creatinine determination based on the Jaffe reaction involved mixing 0.1 mL of sample with 2.5 mL alkaline picrate reagent, incubating for 10 min at room temperature, then measuring the absorbance at 530 nm in a 1-cm cuvette in a spectrophotometer set to read zero using a cuvette containing distilled water. The following readings were obtained:

Blank (water as ample)	0.050
Creatinine standard (200 µmol/L)	0.250
Serum sample	0.125
Urine sample (prediluted 1 in 50 with water)	0.200

    Calculate the creatinine concentration in the serum (in µmol/L) and urine (in mmol/L).

11. A standard curve for a plasma glucose method was set up by preparing a series of dilutions of a stock glucose standard (containing 50 mmol glucose/L) and measuring the absorbance at 500 nm in a 1 cm cuvette using a blank with zero glucose concentration to zero the instrument. The following readings were obtained:

Glucose (mmol.L):	5	10	15	20	25	30
Absorbance:	0.102	0.203	0.305	0.375	0.410	0.432

    Does the method obey Beer's Law? What glucose concentration corresponds to an absorbance reading of 0.250?

Chapter 5

Renal function

The kidney has multiple functions but in routine clinical practice very few of these are formally assessed. A proportion of the blood supplied to each kidney is filtered at the *glomerulus* to produce a cell-free *ultrafiltrate* which is virtually protein-free but otherwise has the same composition as plasma. The *tubules* then modify this filtrate by reclaiming components (*reabsorption*) or by adding further components to it (*secretion*) before it is transported via the ureters to the bladder then excreted. In other words the tubules convert the *glomerular filtrate* into *urine*. These two processes, filtration or *glomerular function* and *tubular function* can be quantitatively assessed in the laboratory.

The Glomerular Filtration Rate (*GFR*)

Glomerular filtration is usually quantified as the rate of formation of filtrate. For a normal adult the *glomerular filtration rate (GFR)* is in the order of 100 mL/min (or 0.10 L/min). In other words, each minute the body produces approximately 100 mL of glomerular filtrate i.e. filters about 100 mL of *plasma*. Using this information it is easy to calculate the total amount of filtrate produced in any given time period e.g. 24 h. If the concentration of any component of plasma freely filtered at the glomerulus is known, then, since the concentration of that component is the same in the filtrate, it is also possible to calculate the total amount of that component filtered in any time period.

Question Q5(1)

A normal subject has a *GFR* of 120 mL/min and a plasma creatinine concentration of 100 µmol/L. Calculate total volume (in litres) of filtrate produced over a 24 h period and the 24 h excretion of creatinine (in mmol) assuming that none of the filtered creatinine is reabsorbed by the tubules.

Answer Q5(1)

The longer the time period the more filtrate is produced. Since 120 mL is produced per minute then twice this amount is produced in 2 minutes (2 x 120 mL = 240 mL), three times this amount in 3 minutes (3 x 120 mL = 360 mL) etc. Therefore if the rate of filtration (the *GFR*) is multiplied by the time period (using the same units of time) then the result is the total volume of filtrate produced over that time period:

Vol filtrate (mL) = *GFR* (mL/min) x time period (min)

In this instance the *GFR* is 120 mL/min and the time period is 24 h. Obviously the same units for time must be used so 24 h is multiplied by number of minutes in an hour (60) to give 24 x 60 = 1440 min. From this the volume of filtrate formed in 24 h can be calculated:

Vol filtrate (mL) = 120 x 1440 = 172800 mL/24 h

To convert from mL to L divide by 1000 (since there are 1000 mL in a L):

Vol filtrate (L) = 172800/1000 = **173 L** (to 3 significant figures)

This volume is considerably greater than the total amount of plasma in the human body (approx 3.5 L) or of the total water content (approx 42 L) and emphasises the importance of the renal tubules in reclaiming the vast majority of filtered water to reduce the volume of filtrate to the daily output of urine (in the order of 1-2 L) and thus avoid dehydration.

Calculation of the total amount of a solute in this volume of filtrate is analogous to calculating the amount of a compound needed to prepare a given amount of solution of a given concentration (see chapter 2). All that is need is to multiply the volume of the solution (in this case glomerular filtrate formed in 24 h) by the concentration of the solute (in this case the concentration of creatinine in the glomerular filtrate will be the same as its plasma concentration):

Creatinine filtered (μmol/24 h) = Volume filtrate (L/24 h) x Plasma creatinine (μmol/L)

= 173 x 100 = 17300 μmol/24 h

Division by 1000 converts to mmol/L (since there are 1000 μmol in 1 mmol):

Creatinine filtered = 17300/1000 = **17.3 mmol/24 h**

If none of this filtered creatinine is reabsorbed by the tubules then this represents the 24 urinary excretion of creatinine.

The manipulations involved in calculating amounts of substances filtered can be summarised:

Volume of filtrate = *GFR* x Time ……..… Eq. 5.1

Amount filtered = Volume of filtrate x Plasma concentration ….. Eq. 5.2

Amount filtered = *GFR* x Time x Plasma concentration ……... Eq. 5.3

Rate of filtration of substance = *GFR* x Plasma concentration ……... Eq. 5.4

It is vital to always ensure that the units are compatible. These calculations always assume that the plasma concentration of the substance remains *constant* over the time period being considered and that it is *freely filtered* at the glomerulus.

Rate of urinary excretion

This is usually assessed by making a timed collection of urine usually, but not always, over a 24 h period. The total volume of urine is measured and the concentration of the analyte of interest measured in an aliquot of the urine collection. Multiplication of the urine volume by the analyte's concentration in the urine aliquot gives its total urinary excretion over the collection period:

Total excreted = Urine volume x Urine concentration ……. Eq. 5.5

Again, care must be taken that units are compatible. Note that if *rate* of urine production is used instead of urine volume then this calculation gives the *rate* of urinary excretion of the compound.

Question Q5.2

A patient was asked to collect urine over a 24 h period. The volume was found to be 1.5 L and the concentration of creatinine in an aliquot of this urine 8.0 mmol/L. Calculate the 24 h urinary excretion of creatinine in mmol.

Answer Q 5(2)

Creatinine excretion(mmol) = Urine volume (L) x Creatinine concentration (mmol/L)

= 1.5 x 8.0 = **12.0 mmol**

Clearance and *GFR*

Consider a hypothetical compound which is only excreted by the kidney (and is not further metabolised by the body) and is in a *steady state* i.e. the rate of excretion via the kidneys is equal to the rate of formation by the body (or rate of infusion into the body) so that its *plasma concentration remains constant*. In practice this can be achieved in two ways:

- Administration of an exogenous compound (e.g. inulin – a fructose polymer) intravenously. By definition when a steady state is reached the rate of excretion is equal to the rate of infusion and urinary collections are unnecessary.

- Using a compound which is produced endogenously by the body at a constant rate (e.g. creatinine).

Provided this compound is freely filtered at the glomerulus then its rate of filtration is given by Eq.5.4:

Rate of filtration = *GFR* x Plasma concentration

And its rate of excretion in the urine can be described by Eq.5.5:

Rate of excretion = Rate of urine formation x Urine concentration

If all of the compound that is filtered at the glomerulus is excreted in the urine i.e. it is neither reabsorbed from nor further amounts secreted into the filtrate, then:

Rate of filtration = Rate of excretion

If we substitute for rates of filtration and excretion the following expression is obtained:

GFR x Plasma concentration = Rate of urine formation x Urine concentration

which can be rearranged to give:

$$GFR = \frac{\text{Rate of urine formation} \times \text{Urine concentration}}{\text{Plasma concentration}} \quad \text{…….. Eq. 5.6}$$

Therefore, provided the above conditions are met, if a timed urine is collected and the concentration of the analyte is measured in *both* plasma and urine, then the *GFR* is easily calculated. Again it is vital that all the *units are compatible*.

Another way of looking at the *GFR* is that it is the clearance of the substance being considered i.e. the volume of plasma from which the substance is completely removed or cleared in unit time. This volume can easily be determined if urinary excretion is divided by plasma concentration:

$$\text{Urinary clearance (L)} = \frac{\text{Urinary excretion}}{\text{Plasma concentration}}$$

If the substance is reabsorbed or secreted by the tubules then of course the measured urinary excretion of the compound will *not* be equal to the rate of its filtration at the glomerulus. In other words the *GFR* will *not* be equal to the measured clearance. If a compound is *reabsorbed* by the tubules (for example urea) then only a proportion of the filtered compound will be excreted in the urine and the measured clearance will be much *lower* than the GFR. In the case of urea, the amount reabsorbed is very dependent upon the state of hydration (and hence the urine flow rate) and attempts have been made to correct for this using the square root of the urine volume in the calculation. If a substance is actually *secreted* into the tubules then the urinary excretion will be greater than the amount filtered and the measured clearance will be *greater* than the *GFR*.

The clearance of any substance can be measured, but the value obtained will only give a measure of *GFR* if it is excreted by glomerular filtration alone.

Question Q 5(3)

A 24 h urine (volume 2.4 L) was collected from a patient in ITU and the creatinine concentration of an aliquot was found to be 6.0 mmol/L. The creatinine concentration of a plasma sample collected during this 24 h period is 500 µmol/L. Calculate this patient's creatinine clearance in mL/min.

Answer Q 5(3)

$$\text{Creatinine clearance} = \frac{\text{Rate of urinary creatinine excretion}}{\text{Plasma creatinine concentration}}$$

$$\text{Urine flow rate (L/min)} = \frac{\text{Urine volume (L/24 h)}}{24 \times 60} = \frac{2.4}{1440} = 0.00167 \text{ L/min}$$

(Division by 24 converts the urine flow to L/h, then division by 60 converts to L/min).

Rate of excretion (µmol/min) = Urine flow rate (L/min) x Urine concentration (µmol/L)

Since the urine creatinine concentration is given in mmol/L it is first multiplied by 1000 to covert to µmol/L, then:

Rate of excretion (µmol/min) = 0.00167 x 6.0 x 1000 = 10.0 µmol/min

Division of this rate of urinary excretion of creatinine by its plasma concentration gives the volume of plasma completely cleared of creatinine in each minute . i.e. the clearance:

$$\text{Clearance (L/min)} = \frac{\text{Rate of excretion (µmol/min)}}{\text{Plasma concentration (µmol/L)}}$$

$$\text{Clearance} = \frac{10.0}{500} = 0.020 \text{ L/min}$$

Multiplication by 1000 converts the clearance from L/min to mL/min:

Clearance = 0.020 x 1000 = **20 mL/min**

Alternatively the equation described in Fig 5.1 could be used.

It should be noted that the calculation of clearance involves three measurements: urine volume, urine concentration and plasma concentration. The combination of the errors involved in these measurements is considerable. It cannot be emphasised too strongly that the *largest source of error in a clearance measurement is the accuracy of the timed urine collection.* Although clearance measurements are conventionally expressed as mL/min, the result implies an unrealistic degree of accuracy and it would be better if results were expressed as L/min using only 2 significant figures. For example a clearance of 123 mL/min would become 0.12 L/min.

The clearance of a substance is given by the equation:

$$\textbf{Clearance} = \frac{U \times V}{P} \quad \text{..... Eq. 5.7}$$

Where: U = concentration of the substance in urine
V = rate of urine formation
P = concentration of the same substance in plasma

The units must *always* be compatible. In the case of creatinine clearance, the clearance is usually expressed as mL/min so the units of the individual measurements are first adjusted as follows:

U: urine creatinine is usually reported in **mmol/L**. Since plasma creatinine is reported in µmol/L, the urine creatinine is multiplied by 1000 to convert it to the same units.

V: urine is usually collected over a **24 h** period and its volume expressed in **litres**. Since the clearance is required in mL/min this volume is multiplied by 1000 (to convert from litres to mL) and divided by 24 (to convert from 24 h to 1 h) then 60 (to convert from hours to minutes).

P: plasma creatinine is usually expressed as **µmol/L** and is unchanged.

Introducing these adjustments leads to the equation:

$$\text{Creatinine clearance (mL/min)} = \frac{U \times 1000 \times V \times 1000}{24 \times 60 \times P}$$

Which (to 2 significant figures) simplifies to:

$$\textbf{Creatinine clearance (mL/min)} = \frac{\boldsymbol{U \times V \times 700}}{P} \quad \text{..... Eq. 5.8}$$

If different units or urine collection times are used then the factor 700 must be adjusted.

Figure 5.1 Calculation of creatinine clearance

Relationship between clearance, plasma concentration and urinary excretion

Equation 5.7, clearance = $\frac{UV}{P}$, contains 3 variables:

Clearance, which may or may not be equal to the *GFR*
Plasma concentration (*P*)
Urinary excretion of the substance (*UV*)

Alteration of one variable must result in the alteration of at least one other. Consider a steady state in which clearance, *P* and *UV* are constant (Fig 5.2).

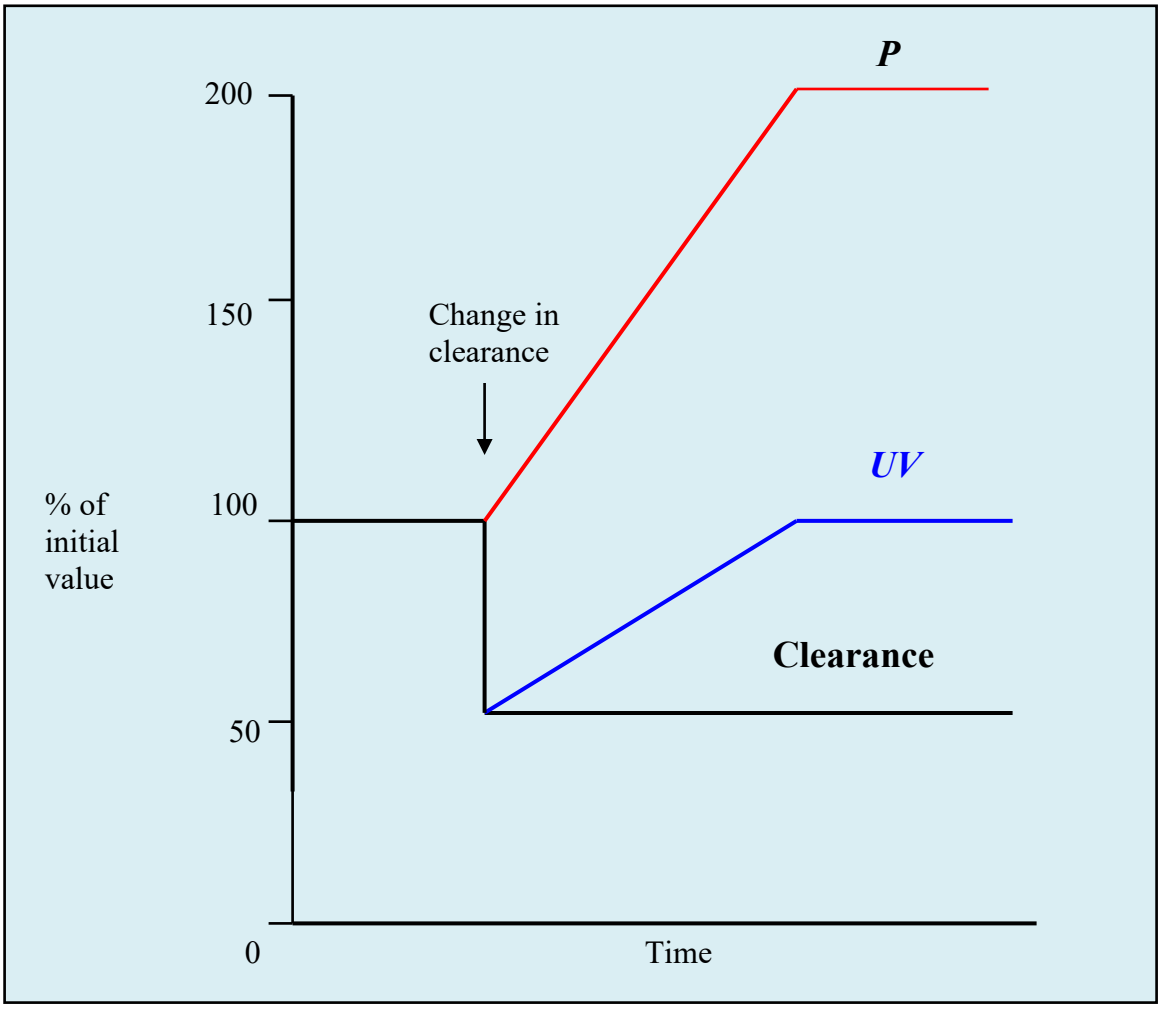


Figure 5.2 Effect of a change in clearance (shown by arrow) on plasma concentration (*P*) and urinary excretion (*UV*) of a substance

If at some point (shown by the arrow) the clearance is *halved* (for example by removal of one kidney) then the immediate effect will be for the urinary excretion (*UV*) to halve. As time progresses the plasma concentration (*P*) will rise (since less of the substance is being

removed from the plasma by filtration) and as a result the amount filtered per unit time (given by: clearance x *P*) will also rise. This will be reflected in the urinary excretion (*UV*) which will also increase. Eventually a new steady state will be achieved in which the urinary excretion is unchanged but the plasma concentration is doubled (Fig 5.2). The relationship between clearance, *P* and *UV* is summarised in Fig 5.3:

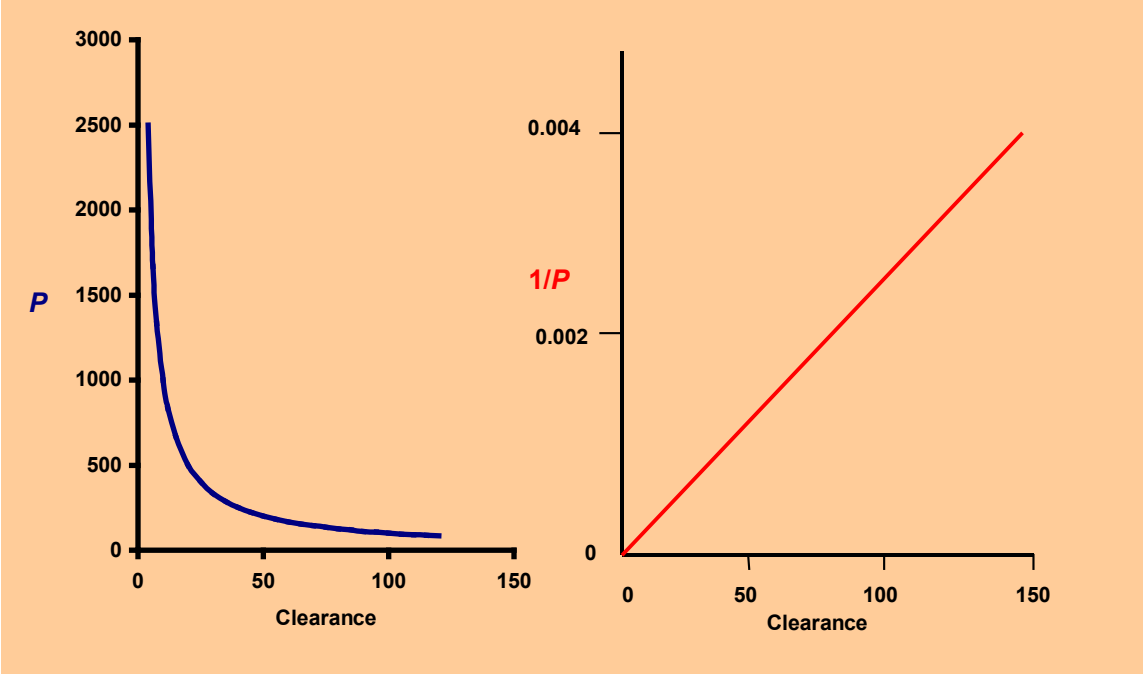


Figure 5.3 Relationship between plasma concentration (*P*), reciprocal of plasma concentration (1/*P*) and clearance

Question Q 5(4)

A patient with a creatinine clearance of 120 mL/min has a plasma creatinine concentration of 100 µmol/L. Assuming no tubular action on filtered creatinine, what concentration of creatinine would you expect to find in a 6h urine collection which has a total volume of 500 mL? A healthy kidney is removed by surgery for transplantation to a relative. What concentration of creatinine would you expect to find in a 6 h urine (volume 400 mL) collected immediately following surgery. At a follow-up clinic six weeks later his plasma creatinine was found to be 200 µmol/L. Estimate the likely creatinine content of a repeat 6 h urine collected the day before his clinic appointment assuming his creatinine clearance is unchanged.

Answer Q 5(4)

$$\text{Clearance (mL/min)} = \frac{(U \times 1000) \times (V \times 1000)}{(h \times 60) \times P} \quad \ldots\ldots \text{Eq. 5.8}$$

$$\text{Rearranging:} \quad U = \frac{\text{Clearance} \times (h \times 60) \times P}{1000 \times (V \times 1000)}$$

$$\text{Clearance} = \frac{U \times V}{P} \ldots\ldots\ldots\ldots\ldots\ldots \text{Eq. 5.7}$$

Pre-operatively:

Clearance = 120 mL/min
h = urine collection period = 6 h
V = urine volume = 500 mL = 0.5 L
P = plasma creatinine = 100 µmol/L

$$U = \frac{120 \times 5 \times 60 \times 100}{1000 \times 0.5 \times 1000} = \textbf{8.6 mmol/L}$$

Immediately following removal of one kidney:
Assuming both kidneys function equally and there is no pre-renal impairment post-operatively then the clearance will be a half of the previous value.

Creatinine clearance = 120/2 = 60 mL/min
P is initially unchanged, *V* = 400 mL and *h* is again 6

Substitution of these values into the rearranged Eq. 5.8 allows calculation of the new value of *U*:

$$U = \frac{60 \times 6 \times 60 \times 100}{1000 \times 0.4 \times 1000} = \textbf{5.4 mmol/L}$$

Follow-up at six weeks:

GFR = unchanged = 60 mL/min
P = 200 µmol/L
Both *V* and *U* are unknown

Eq. 5.8 can be rearranged in a slightly different way to give:

$$U \times V = \frac{\text{Clearance} \times (h \times 60) \times P}{1000 \times 1000}$$

Therefore $U \times V = \frac{60 \times 6 \times 60 \times 200}{1000 \times 1000} =$ **4.3 mmol**

Correction of *GFR* for Body Surface Area

Kidney size and hence GFR increases with increasing body size. This is not surprising since as body size increases so does body water and hence plasma volume. Furthermore, a larger body produces a greater amount of waste products to be excreted. For example, the GFR of a 2 year-old child is about 45 mL/min, a reflection of lower body size than an adult.

To compensate for variations in body size, GFR or clearance is often related to body surface area. Therefore results are either expressed as mL/min/m^2 or relative to an "average" body surface area of 1.73 m^2 i.e. as mL/min/1.73 m^2. To do this it is first necessary to calculate the body surface area. Body surface area can be estimated from both height (in cm) and body weight (in Kg) using the following formula:

$$\log_{10} A = (0.425 \times \log_{10} W) + (0.725 \times \log_{10} H) - 2.144$$

where A is body surface area in m^2 , W is body weight (in Kg) and H is height (in cm).

If the measured *GFR* or clearance is divided by this surface area then the result becomes mL/min/m^2. If this result is then multiplied by 1.73 then the *GFR* or clearance is corrected to the average or standard surface area of 1.73 m^2. These calculations are summarised in Fig 5.6.

Question Q 5(5)

A patients measured creatinine clearance is 60 mL/min. Correct this value to a standard body surface area of 1.73 m^2 given that the body weight is 38.5 kg and height 102 cm.

Correction of *GFR* (or clearance) for body surface area

$$A = \text{antilog}_{10}\,[\,(0.425 \times \log_{10} W) + (0.725 \times \log_{10} H) - 2.144] \quad \text{Eq. 5.10}$$

Where: A = body surface area in m^2
W = body weight in Kg
H = body height in cm

GFR expressed per m^2 body surface area:

$$\text{Corrected } GFR\ (\text{mL/min/m}^2) = \frac{\text{Measured } GFR\ (\text{mL/min})}{A\ (\text{m}^2)} \quad \text{Eq. 5.11}$$

GFR expressed as a ratio to "standard" surface area of 1.73 m^2:

$$\text{Corrected } GFR\ (\text{mL/min/1.73m}^2) = \frac{\text{Measured } GFR\ (\text{mL/min}) \times 1.73}{A\ (\text{m}^2)} \quad \text{Eq. 5.12}$$

Figure 5.6 Correction of *GFR* (or clearance) for body surface area

Again it should be emphasized that this correction is NOT the same as estimating clearance from plasma creatinine measurement. In order to calculate clearance form plasma creatinine then the body weight (W) will be used *twice*: once to estimate clearance (uncorrected for body size) from plasma creatinine (using Eq. 5.9) then again to calculate body surface area (using Eq. 5.11) for correction of this value to standard body surface area.

Answer Q 5(5)

First calculate the surface area using Eq 5.11:

$$A = \text{antilog}_{10}\,[(0.425 \times \log_{10} W) + (0.725 \times \log_{10} H) - 2.144]$$

Substitute $W = 38.5$ kg and $H = 102$ cm then evaluate A:

$$A = \text{antilog}_{10}\,[(0.425 \times \log_{10} 38.5) + (0.725 \times \log_{10} 102) - 2.144]$$

$= \text{antilog}_{10}\,[(0.425 \times 1.59) + (0.725 \times 2.00) - 2.144]$

$= \text{antilog}_{10}\,[0.676 + 1.45 - 2.144]$

$= \text{antilog}_{10} - 0.018 = 0.96\ m^2$

Next correct the clearance for a body surface area of $1.73 m^2$ using Eq. 5.12:

$$\text{Corrected clearance (mL/min/1.73 m}^2\text{)} = \frac{\text{Measured clearance} \times 1.73}{A}$$

Substitute: measured clearance = 60 mL/min and $A = 0.96\ m^2$

$$\text{Corrected clearance} = \frac{60 \times 1.73}{0.96} = \mathbf{108\ mL/min/1.73\ m^2}\ \text{(2 sig figs)}$$

In this example correction of a low clearance for the small body size resulted in a normal value.

Calculation of creatinine clearance directly from plasma creatinine concentration

Since the largest source of error in a clearance measurement is the accuracy of the timed urine collection and continuous urine collections are very inconvenient, attempts have been made to derive the urinary excretion of creatinine from sources other than urinary analysis. Creatinine originates from two sources:

- From creatine (and creatine phosphate) which is released continuously from muscle. Under normal circumstances the rate of creatinine production is relatively constant and proportional to the total body muscle mass.
- From dietary sources (including dietary creatine).

Therefore methods of calculation have been devised which attempt to relate muscle mass to the rate of creatinine added to the body pool (and hence excreted in urine). Dietary sources of creatinine are usually ignored.

a) Method of Cockroft and Gault

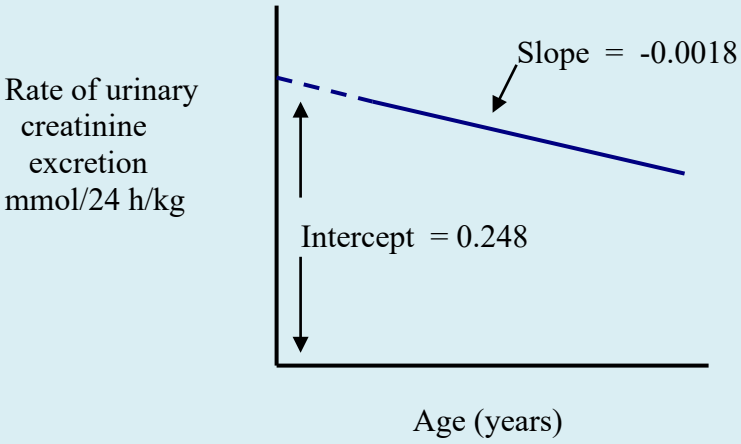
Cockroft and Gault measured urinary creatinine output and derived its relationship with both body weight and age which they then used to develop a formula which estimates

creatinine clearance from plasma creatinine, age and body weight (Fig 5.5). This equation assumes that the relationship between body weight and muscle mass is constant and may be unreliable in obese and oedematous subjects. Note that body weight and age are ONLY used to estimate urinary creatinine output which is then incorporated into the standard equation for creatinine clearance (Eq. 5.8). It does NOT correct the clearance to body surface area – if this is required then an additional calculation is required.

Question Q 5(5)

A 40-year old lady has a plasma creatinine concentration of 153 μmol/L. Estimate her creatinine clearance (in mL/min) given that her body weight is 60 kg.

The Cockcroft-Gault equation



Rate of urinary excretion (mmol/24 h/kg) = (-0.0018 x age) + 0.248

The rate of urinary excretion is the same as U x V.

U x V (mmol/24 h) = Body wt (kg) [0.248 - (0.0018 x age)]

Simplifying:

$$U \times V = 0.0018 \text{ Body wt } \left(\frac{0.248}{0.0018} - \frac{0.0018}{0.0018} \times \text{age}\right)$$

$$U \times V = 0.0018 \text{ Body wt (kg) } (138 - \text{age})$$

U x V is then substituted into Eqn 5.9:

$$\text{Creatinine clearance} = 0.0018 \frac{\text{Body wt (kg) } (138 - \text{age}) \times 700}{\text{Plasma creatinine } (\mu\text{mol/L})}$$

Which simplifies to the following expression (to 2 sig figs):

Creatinine clearance (mL/min) = $\frac{(140 - \text{age (y)}) \times \text{Body wt (kg)} \times 1.3}{\text{Plasma creatinine } (\mu\text{mol/L})}$ …. Eq. 5.13

The above data was validated for men. It is common practice to multiply this value by 0.**85** for women to correct for a lower muscle mass.

Figure 5.5. The Cockroft and Gault equation for predicting creatinine clearance from age, body weight and plasma creatinine

Answer Q 5(6)

$$\text{Creatinine clearance (mL/min)} = \frac{(140 - \text{age in yrs}) \times \text{Body Wt (kg)} \times 1.2}{\text{Plasma creatinine (µmol/L)}}$$

Substitute
Age = 40 years
Body weight = 60 kg
Plasma creatinine = 153 µmol/L

$$\text{Creatinine clearance (mL/min)} = \frac{(140 - 40) \times 60 \times 1.2}{153}$$

$$= \frac{100 \times 60 \times 1.2}{153} = = 47 \text{ mL/min (2 sig figs)}$$

Since the patient is female this result is multiplied by 0.85 to correct for a lower muscle mass than males (for which the formula was derived).

Creatinine clearance = 47 x 0.85 = **40 mL/min** (2 sig figs)

b) Formulas derived from comparative studies of *GFR* and serum creatinine

A purely mathematical approach is employed in order to derive equations which convert serum creatinine to estimated *GFR* (*eGFR*), taking into account variables such as age, sex and ethnicity which are likely to affect creatinine production. *GFR* was always obtained by a reliable clearance method (e.g. iothalamate) and serum creatinine measurements aligned to a reference method. *GFR*, and hence *eGFR*, is expressed for an average body surface area of 1.73 m^2.

The earliest equation to be widely recommended in most official guidelines was derived by the Modification of Diet in Renal Disease (MDRD) study group. Later the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) published a new equation using pooled data from multiple studies for both its development and validation[1]. Their approach was a little different in that they divided the data into two groups – those with serum creatinine below a certain value (called a "knot") and those above it – resulting in two equations (although they can be combined into one provided appropriate constants are selected according to plasma creatinine and sex). The CKD-EPI equation performed better than the MDRD equation, especially at higher GFR. At the time of writing the CKD-EPI equation is recommended by the National Institute of Clinical Excellence (NICE). An easy to use version of this equation is included in Fig 5.6.

[1] Levey AS, Stevens LA, Schmid CH *et al*. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009; 150(9): 604-612.

These formulae are still undergoing development and it is likely that guidelines may change in the near future.

Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) formula

$$\text{GFR (mL/min/1.73 m}^2\text{)} = 141 \times \left[\frac{\text{serum creatinine (µmol/L)} \times 0.011312}{k}\right]^{\alpha} \times 0.993^{\text{Age}} \times [1.018 \text{ if female}] \times [1.159 \text{ if black}]$$

Appropriate values for α and k are:

	Male		Female	
Serum creatinine (µmol/L)	**≤80**	**>80**	**≤62**	**>62**
k	**0.9**	**0.9**	**0.7**	**0.7**
α	**-0.411**	**-1.209**	**-0.329**	**-1.209**

Fig 5.6 CKD-EPI formula for estimation of *GFR* (*eGFR*) from serum creatinine concentration

Question Q 5(7)

Use the CKD-EPI formula to estimate the GFR for a 55-year old Caucasian woman whose serum creatinine is 125 µmol/L.

Answer Q 5(6)

Substitute into the CKD-EPI equation:

Serum creatinine = 125 µL Age = 55 years
k = 0.7 (since patient is female)
α = -1.209 (since patient is female with creatinine > 62)
Include the factor of 1.018 since the patient is female but not the factor of 1.159 since she is Caucasian.

$$GFR = 141 \times \left[\frac{125 \times 0.011312}{0.7}\right]^{-1.209} \times 0.993^{55} \times 1.018$$

$$= 141 \times 0.4274 \times 0.6795 \times 1.018 = \mathbf{42\ \mu mol/L/1.73\ m^2}$$

Measures of Tubular Function – Fractional Excretion (*FE*) and Tubular Maximum (*Tm*)

The fractional excretion (*FE*) expresses the excretion of a solute as a fraction of the amount filtered at the glomeruli:

$$\text{Fractional excretion } (FE) = \frac{\text{Amount excreted in urine}}{\text{Amount filtered}}$$

$$\text{or } \frac{\text{Rate of urinary excretion}}{\text{Rate of filtration}}$$

The rate the substance is excreted in urine can be calculated from the concentration of the substance in urine ($U_{substance}$) and the urine flow rate (V):

$$\text{Rate of excretion} = U_{substance} \times V$$

The rate of filtration is the product of *GFR* and plasma concentration (P):

$$\text{Rate of filtration} = P_{substance} \times GFR$$

These two expressions can be combined to give a formula to calculate $FE_{substance}$:

$$FE_{substance} = \frac{U_{substance} \times V}{P_{substance} \times GFR} \quad \text{.....Eq. 5.14}$$

If *GFR* is estimated at the same time and using the same urine collection by measuring the concentrations of creatinine in the same plasma and urine, then:

$$GFR = \frac{(U_{creatinine} \times V)}{P_{creatinine}}$$

This value for *GFR* can then be substituted into Eq. 5.14 to give:

$$FE_{substance} = \frac{U_{substance} \times V \times P_{creatinine}}{P_{substance} \times U_{creatinine} \times V}$$

The *V* terms cancel so that the expression simplifies to:

$$\boldsymbol{FE_{substance} = \frac{U_{substance} \times P_{creatinine}}{U_{creatinine} \times P_{substance}}} \quad \textbf{.....Eq. 5.15}$$

Question Q 5.8

The following results were obtained in a 20 year-old male admitted after a car crash and found to be oliguric:

Plasma: Sodium = 125 mmol/L Creatinine = 200 µmol/L

Urine: Sodium = 60 mmol/L Creatinine = 1.2 mmol/L

Calculate the fractional excretion of sodium.

Answer Q5(8)

$$FE_{Na} = \frac{U_{Na} \times P_{creat}}{U_{creat} \times P_{Na}}$$

Substituting the relevant values after making sure that units are appropriate:

P_{Na} = 125 mmol/L
U_{Na} = 60 mmol/L
P_{creat} = 200 µmol/L
U_{creat} = 1.2 mmol/L = 1.2 x 1000 = 1200 µmol/L

$$FE_{Na} = \frac{60 \times 200}{1200 \times 125} = \mathbf{0.08}$$

Many substances filtered at the glomerulus are reabsorbed by the tubules. If the rate of filtration is less than the rate of tubular re-absorption for the substance then the urinary excretion is *zero*. The *proportion* of the filtered substance that is reabsorbed is called the *tubular reabsorption* (*TR*). Since the rate of filtration of a substance is given by the product of its concentration and *GFR*, the rate of filtration increases with plasma concentration. Most tubular transport mechanisms are concentration dependent so as the concentration of the substance in the glomerular filtrate increases the rate of reabsorption also increases but eventually a threshold is reached as the tubular transport mechanism becomes saturated and no further substance can be reabsorbed and any excess appears in the urine. This maximal rate of tubular transport is called the *Tm* value of the substance. This gives rise to the concept of the *renal threshold* for a substance and is illustrated in Fig 5.7. Since by definition the proportion filtered is 1, the numerical relationship between *FE* and *TR* is:

$$FE = 1 - TR \quad \text{..... Eq. 5.16}$$

If the rate of filtration (*GFR* x *P*) is less then the value for *TR* then none of the substance appears in the urine and the value for *FE* is negative. However, if the rate of filtration exceeds the *TR* value then the renal threshold is exceeded and the substance appears in urine and is possible to calculate *TR* by rearranging Eq. 5.17:

$$TR = 1 - FE \quad \text{..... Eq. 5.17}$$

In other words the fraction that is not excreted must have been reabsorbed. To convert this fraction to an absolute concentration (i.e. the amount of substance reabsorbed per given volume of filtrate) then the value for *TR* is multiplied by the plasma concentration. This is the same thing as the ratio of the maximal rate of reabsorption (*Tm*) to the GFR – i.e. the *Tm/GFR*:

$$Tm/GFR = TR \times P \quad \text{..... Eq. 5.18}$$

This method of calculation assumes that the plasma concentration (*P*) is well above the renal threshold. Units are concentration e.g. mmol/L.

Question Q 5(9)

The following results were obtained for urine and plasma from a fasting adult:

Plasma phosphate	0.65 mmol/L	Plasma creatinine	105 µmol/L
Urine phosphate	11.5 mmol/L	Urine creatinine	4.64 mmol/L

Calculate: a) The fractional excretion of phosphate (*FEP*)
b) The fractional tubular reabsorption of phosphate (*TRP*)
c) The renal tubular reabsorption of phosphate (*TmP/GFR*)

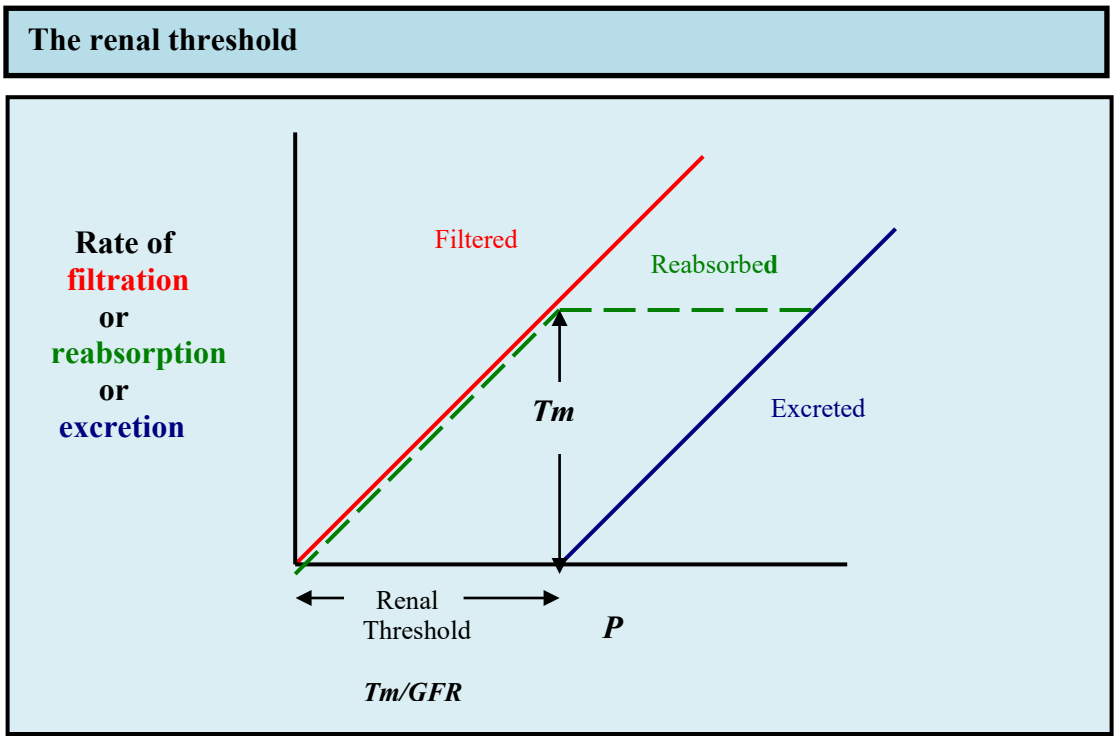


Figure 5.7 The renal threshold for a substance filtered at the glomerulus then reabsorbed by the tubules

Answer Q5(9)

a) To calculate the fractional excretion of phosphate (*FEP*) use Eq. 5.15:

$$FEP = \frac{U_{phosphate} \times P_{creatinine}}{U_{creatinine} \times P_{phosphate}}$$

Where:
$U_{phosphate}$ = 11.5 mmol/L
$P_{creatinine}$ = 105 µmol/L
$U_{creatinine}$ = 4.64 mmol/L = 4.64 x 1000 = 4640 µmol/L
$P_{phosphate}$ = 0.65 mol/L

Substituting these values:

$$FEP = \frac{11.5 \times 105}{4640 \times 0.65} = \mathbf{0.40}$$

b) The proportion of filtered phosphate that is reabsorbed (*TRP*) is the difference between the fraction filtered (by definition 1) and the fraction excreted (*FEP*):

$$TRP = 1 - FEP$$

$$TRP = 1 - 0.40 = \mathbf{0.60}$$

c) Since the *TRP* is the fraction of filtered phosphate that is reabsorbed, multiplication by the plasma phosphate concentration gives the maximum rate of reabsorption of phosphate (in mmol) per litre of glomerular filtrate (*TmP*/*GFR*):

$$TmP/GFR = TRP \times P_{phosphate}$$

$$TmP/GFR = 0.60 \times 0.65 = \mathbf{0.39\ mmol/L\ glomerular\ filtrate}$$

Reference: Payne RB. Renal tubular reabsorption of phosphate (*TMP*/*GFR*): indications and interpretation. *Ann Clin Biochem* 1998; **35**: 201-206.

Osmolar Clearance and Free Water Clearance

If osmolality is substituted for creatinine in the calculation of creatinine clearance then the result (Fig 5.8) is the *osmolar clearance* (C_{osm}). Osmolality is a measure of the sum of the individual osmotically active species present in plasma or urine (see chapter 6). Just as the creatinine clearance can be thought of as the volume of plasma from which all creatinine is removed per unit of time, osmolar clearance can be regarded as the volume of plasma from which all filterable solutes are removed in any given time period. It can also be thought of as the minimum volume of urine required to excrete a solute load in *isosmolar* (the same osmolality as plasma) form.

The calculated difference between the actual urine volume and the osmolar clearance is known as *free water clearance*. In a hyposmolar urine the free water clearance is positive; in a hyperosmolar urine it is negative. The free water clearance can also be viewed as the volume of pure water subtracted from (positive free water clearance) or added to (negative free water clearance) the plasma per unit time.

The renal excretion of a large water load is limited by the diluting capacity of the tubules and the amount of solute available for excretion. The minimum osmolality of urine is about 50 mmol/kg. On a medium protein diet 1200 mmol is excreted per day (mainly urea) and the maximum urine volume is 1200/50 = 24 L. In starvation however, the source of solute is tissue breakdown (approx. 100-200 mmol/day. There is little ability to excrete free water and hyponatraemia myoccur if water intake is greater than 2 to 4 L/day in starvation.

Calculation of osmolar clearance and free water clearance

If: V = urine flow rate
P_{osm} = plasma osmolality
U_{osm} = urine osmolality

Then the *osmolar clearance* (C_{osm}) is given by:

$$C_{osm} = \frac{U_{osm} \times V}{P_{osm}} \quad \text{..... Eq. 5.19}$$

And the *free water clearance* (C_{water}) is given by:

$$C_{water} = V - C_{osm} \quad \text{..... Eq. 5.20}$$

Figure 5.8 Calculation of osmolar clearance and free water clearance

Question Q 5(10)

A 24 h urine collection (volume 1.20 L) has an osmolality of 750 mOsm/kg. If the plasma osmolality is 300 mOsm/kg calculate:

a) The osmolar clearance in mL/min
b) The free water clearance in mL/min.

Answer Q 5(10)

a) Using Eq. 5.19: $C_{osm} = \frac{U_{osm} \times V}{P_{osm}}$

P_{osm} = 300 mOsm/Kg
U_{osm} = 750 mOsm/Kg
V = urine flow rate = 1.20 L/24 h

Since *Cosm* is required in mL/min, the flow rate must first be converted to mL/min by multiplying by 1000 (1000 mL in 1L) and dividing by 24 (to convert to h) then 60 (to covert to min):

$$V = \frac{1.20 \times 1000}{24 \times 60} = 0.833 \text{ mL/min}$$

Substituting into Eq. 5.19:

$$C_{osm} = \frac{750 \times 0.833}{300} = \mathbf{2.08\ mL/min}$$

b) Calculate the free water clearance using Eq 5.20

$$C_{water} = V - C_{osm}$$

Ensuring that the units are the same:

V = 0.833 mL/min
C_{osm} = 2.08 mL/min

C_{water} = 0.833 - 2.08 = **-1.25 mL/min** (2 sig figs)

FURTHER QUESTIONS

1. An aliquot of a 24 h urine (volume 1850 mL) has a creatinine concentration of 8500 µmol/L. Calculate the 24 h urinary creatinine excretion expressing the result as mmol/24 h.

2. A patient has a *GFR* of 110 mL/min. If the plasma creatinine concentration is 180 µmol/L how many mmol of creatinine are filtered in 12 h?

3. A urine collection (volume 3.2 L) was handed in by a patient which he said he had collected over the previous day. Calculate the creatinine clearance given that the urine was found to have a creatinine concentration of 7.2 mmol/L. The plasma creatinine concentration taken during the collection was 94 µmol/L. Give the most likely cause for this result.

4. The concentration of a compound in the plasma of a normal adult is 10 mg/L. The *GFR* is 110 mL/min and 316.8 mg of the compound are excreted over 24 h in a urine volume of 1584 mL. Comment on these findings.

5. A subject with a *GFR* of 100 mL/min was infused with a 'drug' X at a rate of 100 µmol/min and the plasma concentration reaches a steady state value of 200 µmol/L. It is known that this drug is not metabolized or excreted by organs other than the kidney. What is the clearance of this drug? Comment on the result.

6. A patient who is severely water depleted and excreted only 100 mL of urine in the last 6 hours was a short time before, found to have a creatinine clearance of 100 mL/min with a plasma creatinine concentration of 100 µmol/L. If renal function has remained unchanged what concentration of creatinine would you expect to find in the latest 100 mL (6 h collection) specimen of urine?

7. Estimate the effect on urinary sodium excretion in a person with a *GFR* of 95 mL/min and plasma sodium of 140 mmol/L, of a 1% decrease in the overall reabsorption of sodium.

8. The following data were obtained for a hypertensive patient on a low sodium diet:

Plasma:	creatinine	= 200 µmol/L	sodium	= 155 mmol/L
24 H Urine:	creatinine	= 12.5 mmol/L	volume	= 1250 mL

If the renal tubules reabsorb 90% of filtered sodium, how many grams of sodium are excreted in the same 24 h period?

9. The following results were obtained in a 20 year old male admitted after a car crash and found to be oliguric:

Plasma	Na	125 mmol/L
	K	5.0 mmol/L
	Urea	25.0 mmol/L
	Creatinine	200 µmol/l
Urine	Na	60 mmol/L
	Creatinine	1.2 mmol/L
	Osmolality	200 mOsm/Kg

Calculate the fractional excretion of sodium.

10. A 45 year old lady has a body weight of 56 kg and a height of 155 cm. If her plasma creatinine is 150 µmol/L estimate her *GFR* expressing the result as mL/min/1.73 m^2.

11. Calculate the tubular maximum reabsorptive capacity (*Tm/GFR*) for glucose from the following data:

Plasma glucose	10 mmol/L	Plasma creatinine	120 µmol/L
Urine glucose	50 mmol/L	Urine creatinine	6.0 mmol/L

The urine (volume 30 mL) was collected over a 15 minute period.

12 A 6 h urine collection (volume 800 mL) has an osmolality of 200 mOsm/kg. If the plasma osmolality is 260 mOsm/kg calculate the free water clearance in mL/min.

13. An estimation of glomerular filtration rate can be calculated using the abbreviated MDRD (Modified Diet in Renal Disease) formula:

GFR (mL/min/1.73 m^2) = 186 x [serum creatinine x 0.011312]$^{-1.154}$ x [age in years]$^{-0.203}$ x 0.742 if female and/or x 1.21 if Afro American origin (where serum creatinine is in µmol/L)

Calculate the GFR for a 55 year old Caucasian women whose serum creatinine is 125 μmol/L, and her creatinine clearance, given that a 24 h urine collection with a volume of 1.1 L had a creatinine concentration of 4.7 mmol/L.

Comment critically on the two values.

Chapter 6

Osmolality

Determination of plasma and urine osmolality can be useful in the assessment of electrolyte disorders. Comparison of plasma and urine osmolalities is invaluable in investigating renal water regulation in the setting of severe electrolyte disturbances as may occur in diabetes insipidus and the syndrome of inappropriate antidiuresis.

What is osmolality?

Osmosis is the process by which *solvent* moves from an area of *low solute* concentration, through a *semi-permeable membrane,* to an area of *high solute* concentration. A semi-permeable membrane is one which is permeable to solvent but not solute. The pressure of the solvent movement will depend on the gradient in solute concentrations separated by the membrane and is known as *osmotic pressure*. In the human body there is no active mechanism for the transport of water into and out of cells; water always follows an osmotic gradient. Most cell membranes are freely permeable to water. The main exception to this are the membranes of cells lining the collecting tubules in the kidney, which only become permeable to water in the presence of antidiuretic hormone (ADH). The cell membranes are however, selectively permeable to solutes such as sodium and glucose a property which is modulated by hormonal action e.g. aldosterone and insulin.

The osmotic pressure of a solution depends upon the concentration of particles in solution, which is termed its *osmolality*. The osmolality of a solution is the total concentration (in mols) of all osmotically active species present. A molecule such as sodium chloride which can dissociate into two ionic particles (sodium and chloride ions) will have an osmolality of approximately twice that of an equimolar solution of a species such as glucose which does not dissociate into constituent ions. However, since electrolyte dissociation is always incomplete and there are always associations between

solute and solvent, electrolytes do not behave in an *ideal* manner. In these circumstances osmolality can be calculated in the following manner:

Osmolality = osmol/kg water = *Ó n C* ..... Eq. 6.1

Where: *Ó* = osmotic coefficient
n = number of particles into which each molecule in the solution potentially dissociates
C = molality in mol/kg water

In routine clinical biochemistry it is usual to assume that osmotic coefficients are always equal to 1.

Confusion often arises between the terms osmol*ality* and osmol*arity*. These terms may be defined as follows:

- Osmol*ality* expresses concentrations relative to the *mass* of *solvent* e.g. mols per *kg of water*.

- Osmol*arity* expresses concentrations relative to volume of solution e.g, mols per *litre* of *solution*.

Osmometers always measure osmolality. Osmotic concentration calculated from molar concentrations gives osmolarity. In routine clinical biochemistry the difference is nearly always ignored and the two terms used interchangeably.

Properties other than osmotic pressure are also dependent upon the concentration of solute particles in solution. These include increasing vapour pressure, raising boiling point and decreasing freezing point. These are all known as *colligative properties* and each may be used to measure to obtain a measure of the active solute concentration of a solution. In routine clinical practice osmometers based on depression of freezing point are usually used to obtain a measure of osmolality.

Question Q 6(1)

Estimate the osmolality (in mOsmol/Kg) of (a) 20% glucose, (b) Physiological saline, (c) a mixture containing equal volumes of 20% glucose and physiological saline, and (d) 50 mmol/L calcium chloride. Assume ideal behaviour.

Answer Q6(1)

(a) Glucose does not dissociate into ions in solution so its osmolality will be approximately equal to its molar concentration.

$$\text{Concentration (mmol/L)} = \frac{\text{Concentration (mg/L)}}{\text{MW}}$$

$$\text{Concentration (mg/L)} = \text{Concentration (g/100 mL)} \times 10 \times 1000$$

Substitute the concentration of 20%, which is the same as 20 g/100 mL:

$$\text{Concentration (mg/L)} = 20 \times 10 \times 1000 = 200{,}000 \text{ mg/L}$$

Glucose formula: $C_6H_{12}O_6$

$C_6 = 6 \times 12 = 72$
$H_{12} = 12 \times 1 = 12$
$O_6 = 6 \times 16 = 96$
$\text{MW} = 180$

$$\text{Concentration (mmol/L)} = \frac{200{,}000}{180} = 1100 \text{ mmol/L (2 sig figs)}$$

Therefore the osmolality of 20% glucose is approximately **1100 mOsm/kg**

(b) Physiological saline contains 0.9% sodium chloride.

$$\text{Concentration (mmol/L)} = \frac{\text{Concentration (mg/L)}}{\text{MW}}$$

$$\text{Concentration (mg/L)} = \text{Concentration (g/100mL)} \times 10 \times 1000$$

Substitute the sodium chloride concentration of 0.9%, which is the same as 0.9g/100 mL

$$\text{Concentration (mg/L)} = 0.9 \times 10 \times 1000 = 9000 \text{ mg/L}$$

Formula sodium chloride : NaCl

Na = 23
Cl = 35.5
MW = 58.5

$$\text{Concentration (mmol/L)} = \frac{9000}{58.5} = 154 \text{ mmol/L}$$

Since each molecule potentially dissociates into two ionic species:

$$NaCl \longrightarrow Na^{+} + Cl^{-}$$

Then assuming ideal behaviour the osmolality will be approximately twice the molar concentration of sodium chloride:

$$\text{Osmolaity} = 2 \times 154 = \mathbf{308\ mOsm/kg}$$

(c) If equal volumes of 20% glucose and physiological saline are mixed then the resulting osmolality will be one half the sum of the individual osmolalities:

$$\text{Osmolality} = \frac{\text{Osmolality}_{\text{20\%glucose}} + \text{Osmolality}_{\text{Physiological saline}}}{2}$$

$$\text{Osmolality} = \frac{1100 + 308}{2} = \frac{1408}{2} = \mathbf{700\ mOsm/kg} \text{ (2 sig figs)}$$

(d) Calcium chloride ($CaCl_2$) potentially dissociates into 3 ionic species:

$$CaCl_2 \longrightarrow Ca^{++} + 2\ Cl^{-}$$

Therefore the osmolality will be approximately three times the concentration:

$$\text{Osmolality} = 3 \times 50 = \mathbf{150\ mOsm/kg}$$

Plasma osmolality and the osmolal gap

The major osmotically active species present in normal plasma are Na^+, Cl^-, glucose and urea. The simplest formula for calculating osmolality from the concentrations of these species is:

Osmolality	=	**2 [Na^+]**	+	**glucose**	+	**urea**	……	**Eq. 6.2**
mOsm/kg		mmol/L		mmol/L		mmol/L		

Another version of this formula includes a term for potassium concentration i.e. 2[K^+]. The concentration of sodium is multiplied by 2 to allow for the associated anions (mainly chloride and bicarbonate). However, this simple formula does not give very good agreement with measured osmolality since the following assumptions are made:

- That all important osmotically active species are accounted for.
- That all potential dissociations are complete.
- That the anions associated with Na^+ and K^+ are free to contribute to omolality and are not part of a macromolecule (e.g. protein).
- That the activity of each species is the same as concentration i.e. the ions exhibit ideal behaviour.
- That the millim*olal* concentration of each ion (mmol/kg *water*) is the same as its millim*olar* concentration (mmol/L *plasma*). This is not true since plasma is approximately 95% water.

In an attempt improve the agreement between measured and calculated osmolality various modifications of the above formula have been proposed. One of the most popular is:

Osmolality	=	**1.86 [Na^+]**	+	**[glucose]**	+	**[urea]**	+	**9**	…….. **Eq. 6.3**
mOsm/kg		mmol/L		mmol/L		mmol/L			

It is important to remember that osmolality is calculated from three individual measurements, each with its own inherent imprecision so that the combined imprecision of the final result may be greater than the imprecision of measured osmolality. Within these limitations agreement between calculated and measured osmolality is reasonable for plasma but NOT for urine.

In routine clinical practice the commonest reason for comparing measured with calculated osmolality is to obtain evidence for the presence of, or to quantify, an unmeasured osmotically active species. If no such species is present then, within the limitations discussed above, there should be good agreement between measured and calculated osmolality. In other words the difference between the two values (known as the *osmotic gap*) should be approximately zero. If the osmolal gap has a significant numerical value then this is good evidence that an unmeasured osmotically active species is present and the size of the gap is proportional to its concentration.

Osmolal gap	=	**Measured osmolality**	-	**Calculated osmolality**	**….. Eq. 6.4**
mOsm/kg		mOsm/kg		mOsm/kg	

The osmolal gap is often calculated if a patient is suspected of having ingested large amounts of a volatile compound such as ethanol, methanol or ethylene glycol. Provided only a single compound has been ingested and its identity is known then it is possible to derive an *approximate* value for its concentration which is adequate to act as a guide for treatment. Apart from the limitations inherent in calculating osmolality, errors may occur because volatile solvents such has ethanol do not behave entirely as expected with some osmometers.

Question Q 6(2)

The following data were obtained on the plasma from an unconscious man:

sodium	=	140	mmol/L
urea	=	7.0	mmol/L
glucose	=	7.0	mmol/L
osmolality	=	353	mosmol/kg

Assuming that the osmolar gap is due solely to ethanol, calculate the plasma ethanol concentration in mg/dL.

Answer Q 6(2)

First calculate the osmolality from the concentrations of sodium, urea and glucose:

$$\text{Calculated osmolality} = 1.86\,[Na^+] + [\text{urea}] + [\text{glucose}] + 9$$

$$= (1.86 \times 140) + 7.0 + 7.0 + 9$$

$$= 283 \text{ mOsmol/kg}$$

The osmolal gap is the difference between measured and calculated osmolality:

$$\text{Osmolal gap} = \text{Measured osmolality} - \text{Calculated osmolality}$$

$$= 353 - 283$$

$$= 70 \text{ mOsmol/kg}$$

If this osmolal gap is due entirely to ethanol, then the ethanol concentration is 70 mmol/L.

To convert to mg/dL, multiply by the MW and divide by 10.

MW ethanol (C_2H_5OH) = (2 x 12) + 16 + (6 x 1) = 46

Therefore, ethanol $= \frac{70 \times 46}{10} =$ **322 mg/dL**

ADDITIONAL QUESTIONS

1. Calculate the approximate osmolality of a glucose/saline infusion containing equal proportions of 5% glucose and 0.9% sodium chloride.

2. Calculate the approximate osmolality of a solution containing 10% mannitol and 0.9% saline (MW mannitol = 182).

3. A patient was mistakenly given 500 mL 20% mannitol ($C_6H_{14}O_6$) intended for the patient in the next bed instead of the same volume of normal (0.9%) saline. Calculate the extra osmolal load given over that which would have resulted from isotonic saline.

4. What increase in plasma osmoality would result from a plasma ethanol concentration of 92 mg/dL?

5. A 45 year old man is brought to casualty following a fit. He had been working alone late in a garage, when he was found by the security guard who called an ambulance. On admission, he has a large bruise on the left temple and is semi-comatose, he smells of alcohol. The admitting team request urea and electrolytes, glucose and an alcohol and blood gas estimation and arrange an urgent CT scan. The results are as follows:

Sodium	141 mmol/L	Urea	3.5 mmol/l
Ethanol	270 mg/dL	Glucose	3.2 mmol/L

The CT scan does not show any bony injury or evidence of intracranial bleed. The neurological registrar is called and asks for an osmolal gap to help provide a quick estimation of whether there is a possibility that other toxic substances present in the garage, such as antifreeze, have been taken in any quantity.

The measured osmolality is 330 mOsm/kg

a) Calculate the osmolal gap
b) Show whether the alcohol concentration explains the observed osmolal gap, explaining any assumptions you make in the process.

Chapter 7

Basic pharmacokinetics

Pharmacokinetics may be defined as what the body does with drugs and includes such processes as absorption, distribution, metabolism and elimination. In practice phamacokinetic factors determine not only the plasma concentration of the "active" drug but the amount of active drug reaching its site of action. *Therapeutic drug monitoring* (TDM) may be defined as the use of drug measurements in body fluids as an aid to the use of drugs in the management (i.e. cure, alleviation or prevention) of disease.

Bio-availability (*F*)

The proportion of administered drug absorbed is termed its bio-availability (*F*) and depends on the drug, the individual and the dosage form of the drug. Bio-availability is defined as the proportion of administered drug which reaches the circulation:

$$\text{Bio-availability } (F) = \frac{\text{Dose reaching circulation}}{\text{Dose administered}} \quad \text{..... Eq. 7.1}$$

This expression can be rearranged to give the dose absorbed for any administered dose:

$$\text{Dose reaching circulation} = F \times \text{Dose administered} \quad \text{..... Eq. 7.2}$$

Salt-conversion factor (*S*)

Often a drug are administered in several chemical forms. For example phenytoin may be given as the free acid (MW 252) or as its sodium salt (MW 274). Clearly the amount of drug administered if the same weight of each compound is given, will be different. A quantity termed the salt-conversion factor or salt fraction (*S*) may be defined:

$$\text{Salt conversion factor } (S) = \frac{\text{Molecular weight of free drug}}{\text{Molecular weight of compound administered}} \quad \text{.. Eq. 7.3}$$

Therefore for sodium phenytoin *S* would be 252/274 = 0.92 equation 7.2 can be modified to incorporate the salt conversion factor:

Dose reaching circulation = *F* x *S* x Dose administered …….. Eq. 7.4

Volume of distribution (V_d)

Once a drug has been absorbed by the body its distribution depends on such factors as its relative solubility in water and fat, binding by plasma protein, availability of active transport mechanisms and regional blood flow. In the laboratory the relationship between the volume of a solution, the weight of chemical used in its preparation and the concentration of the substance is familiar to us all:

$$\text{Concentration} = \frac{\text{Weight}}{\text{Volume}} \quad \textit{and} \quad \text{Volume} = \frac{\text{Weight}}{\text{Concentration}}$$

Therefore for a given weight of a substance its concentration is inversely proportional to volume The same concept can be applied to a drug, the concentration of which is measured in plasma:

$$\textbf{Volume of distribution } (V_d) = \frac{\textbf{Amount of drug in body}}{\textbf{Plasma concentration}} \quad \text{…… Eq. 7.5}$$

The volume of distribution is a theoretical concept. If the drug is distributed throughout the ECF only then V_d will approximate to the ECF volume. If the drug is lipid soluble then V_d may be far in excess of the volume of total body water.

Question Q7(1)

Calculate the theoretical maximum plasma concentration if 500 mg of the sodium salt of a drug is administered to a 70 kg male. Assume the drug (MW of the parent drug 345 Daltons) is only distributed throughout the extra-cellular fluid (the volume of which is 20% of the body weight) and its bio-availability is 0.85.

Answer Q7(1)

First calculate the salt-conversion factor (*S*):

$$S = \frac{\text{MW parent drug}}{\text{MW sodium salt of drug}}$$

MW parent drug (free acid) = 345

For its sodium salt, Na (atomic weight 23) replaces a hydrogen atom (atomic weight 1).

Therefore MW sodium salt of drug = 345 - 1 + 23 = 367

$$S = \frac{345}{367} = 0.94$$

Dose absorbed = F x S x Dose administered

Since F = 0.85 and 500 mg was administered:

Dose absorbed = 0.85 x 0.94 x 500 = 400 mg

The volume of distribution (V_d) is the ECF volume which is 20% of the body weight (70 kg) – assuming density of ECF is approximately 1:

$$V_d = 70 \times \frac{20}{100} = 14 \text{ litres}$$

$$V_d = \frac{\text{Dose absorbed}}{\text{Plasma concentration}}$$

Therefore: $14 \text{ (litres)} = \frac{400 \text{ (mg)}}{\text{Plasma concentration (mg/L)}}$

Re-arranging: Plasma concentration = $\frac{400}{14}$ = **29 mg/L** (2 sig figs)

CLEARANCE OF A DRUG

Drugs can be cleared from the body by two principal routes:

- Renal excretion (Cl_r)
- Metabolism by the liver (Cl_m)

The total clearance of a drug is the sum of both individual clearances:

$$\text{Total clearance} = Cl_r + Cl_m$$

In practice hepatic function is difficult to quantify and most dosage calculations take into account impairment of renal function only.

Following a single dose of a drug a plot of plasma drug concentration versus time is usually non-linear (Fig 7.1). As the drug is cleared the amount of drug removed from plasma in unit time (i.e. the rate of fall in concentration) decreases as plasma concentration falls. In fact the rate of fall with time can be expressed as:

$$\frac{d\,Cp_t}{d\,t} \text{ is proportional to } Cp_t$$

where Cp_t is the concentration of the drug at time t.

The symbols $d\,Cp_t$ does NOT mean d multiplied by Cp_t but a very small (infinitesimally so) change in Cp_t. Similarly d t is a minute change in t. Mathematicians call $d\,Cp_t / d\,t$ the *first differential* of concentration with respect to time and it can be regarded as the slope of a tangent drawn to the curve at any defined point (i.e. any value of t). When the differential $d\,Cp_t / d\,t$ is proportional to a single concentration term then the elimination process is said to follow *first-order* kinetics. If $d\,Cp_t / d\,t$ is independent of concentration (i.e. constant) then the elimination is said to follow *zero-order* kinetics and, if proportional to two concentration terms, *second-order* kinetics etc.

If a proportionality constant (k_d) is introduced then the following equation is obtained:

$$\frac{d\,Cp_t}{d\,t} = k_d . Cp_t$$

where k_d is known as the *elimination rate constant* which has units time.$^{-1}$ This equation can be rearranged to give:

$$\frac{\mathrm{d}\, Cp_t}{Cp_t} = k_d.\, \mathrm{d}\, t$$

In practice it is easier to measure absolute concentrations (and this is what we require in therapeutic calculations) than it is to measure the rates of change in concentration, so mathematicians use a technique known as *integration* to convert the rate equation into an expression relating absolute concentration to time. Integration is always carried out between defined *limits* – in this case between zero time (when the initial concentration is given the symbol Cp_0) and time t (when the concentration is given the symbol Cp_t). Integration of the above expression between these limits gives the equation:

$$\boldsymbol{Cp_t = Cp_0 \times e^{-kd.t}} \quad \textbf{..... Eq. 7.6}$$

Where e is a universal constant which is encountered in nature whenever we are dealing with exponential growth or decay and has a value of approximately 2.718. Eq 7.6 is still non-linear but can be converted into a linear form by taking logarithms. However, since e appears in the equation it is more convenient to take logarithms to the base e rather than to the base 10. Logarithms to base e ($\log_e$ is conventionally denoted ln) are also known as *natural logarithms* or *Napierian logarithms*. Logarithms to the base e behave exactly the same as logarithms to the base 10 (see Fig 3.2). For example $\ln a.b = \ln a + \ln b$, $\ln a/b = \ln a - \ln b$ and $\ln a^b = b.\ln a$. Therefore $\ln e^{-kd.t} = -k_d t.\ln e$ which, since $\ln e = 1$, becomes $-k_d.t.$ Therefore taking natural logarithms of Eq 7.8 gives the expression:

$$\boldsymbol{\ln Cp_t = \ln Cp_0 - k_d.t} \quad \textbf{..... Eq. 7.7}$$

Therefore if $\ln Cp_t$ is plotted against t then a straight line is obtained with slope $-k_d$ and intercept on the vertical axis equal to $\ln Cp_0$ (see Fig 7.1).

If $\log_{10} Cp_t$ is plotted instead of $\ln Cp_t$ then a straight line is still obtained, but this time the slope is equal to $-k_d/2.303$. This value can be derived using the relationship between natural and common logarithms: $\ln x = \ln 10 \times \log_{10} x = 2.303 \log_{10} x$. Eq. 7.7 can then be written:

$$2.303 \log_{10} Cp_t = 2.303 \log_{10} Cp_0 - k_d.t$$

and dividing both sides by 2.303 gives:

$$\log_{10} Cp_t = \log_{10} Cp_0 - \frac{k_d . t}{2.303}$$

Most modern calculators readily produce natural logarithms so it is rarely necessary to work in common logarithms. The elimination rate constant, k_d, is easily calculated if we have two plasma drug concentrations at different times:

If Cp_1 = plasma drug concentration at time t_1
and Cp_2 = plasma drug concentration at time t_2

then each pair of values can be substituted into Eq 7.7 to give the following two equations:

$$\ln Cp_1 = \ln Cp_0 - k_d . t_1$$
$$\ln Cp_2 = \ln Cp_0 - k_d . t_2$$

Subtraction of one equation from the other eliminates the Cp_0 term to give:

$$(\ln Cp_1 - \ln Cp_2) = - k_d . t_1 + k_d . t_2$$

which can be rearranged to give an expression for kd:

$$\boldsymbol{k_d = \frac{(\ln Cp_1 - \ln Cp_2)}{(t_2 - t_1)}} \quad \textbf{…….. Eq. 7.8}$$

The units of k_d are time.$^{-1}$

Question Q 7(2)

A 15 year old boy presents to casualty following a convulsion. It turns out that he had swallowed 30 of his mother's lithium tablets about 10 hours previously. On admission his lithium concentration is 4.1 mmol/L. A decision needs to be made whether to haemodialyse him to reduce the lithium concentration. As this is not going to be available quickly, the physicians want to know how long he will have toxic levels just with endogenous clearance. How long it will be before his lithium concentration drops to the relatively safe level of 1.5 mmol/L below which toxicity is unlikely, given an elimination rate constant of 0.05 h.$^{-1}$

Answer Q 7(2)

Let Cp_0 = initial drug concentration = 4.1 mmol/L
Cp_t = safe drug level at time t = 1.5 mmol/L
t = time taken to achieve the safe level of 1.5 mmol/L = ??
k_d = elimination rate constant = 0.05 h-1

Substitute these values into the first order rate equation (Eq. 7.7) and solve for t:

$$\ln Cp_t = \ln Cp_0 - k_d . t$$

$$\ln 1.5 = \ln 4.1 - 0.05 . t$$

$$0.405 = 1.411 - 0.05 . t$$

$$0.05 . t = 1.411 - 0.405 = 1.006$$

$$t = \frac{1.006}{0.05} = \textbf{20.12 h (20 h to 2 sig figs)}$$

A quantity known as the *elimination half-life* ($t_{1/2}$) can be defined as the time taken for the plasma drug concentration or total body content of the drug to fall by 50%. Fig 7.1 shows that after one half-life the plasma concentration falls to 50%, after two half-lives to 25%, after three half-lives to 12.5% etc.

The elimination rate constant is related to the elimination half-life. If the initial drug concentration is Cp_0, then after one half-life (when $t = t_{½}$), the concentration Cp_t will be one half of the initial value i.e. $Cp_0/2$. If these values are substituted into the logarithmic form of the first-order rate equation (Eq. 7.7) then the following expression is obtained:

$$\ln Cp_0 / 2 = \ln Cp_0 - k_d . t_{1/2}$$

which can be rearranged to give:

$$k_d . t_{½} = \ln Cp_0 - \ln Cp_0 / 2$$

Since the difference between the logarithms of two individual numbers is the same as the logarithm of one number divided by the other, this expression can also be written:

$$k_d . t_{½} = \ln . \frac{Cp_0 \times 2}{Cp_0}$$

The Cp_0 terms cancel so this expression becomes:

$$k_d . t_{½} = \ln 2$$

ln 2 is 0.693, so rearrangement gives the following expression for the half-life:

$$\boldsymbol{t_{½} = \frac{0.693}{k_d}} \quad \textbf{..... Eq. 7.9}$$

Rearranging Eq 7.9 gives $k_d = 0.693 / t_{½}$ which can then be substituted into the Eq. 7.7 to give an alternative formula:

$$\ln Cp_t = \ln Cp_0 - \frac{0.693 . t}{t_{1/2}}$$

which can be rearranged to give the following alternative expression for half-life:

$$\boldsymbol{t_{½} = \frac{0.693 . t}{\ln Cp_t - \ln Cp_0}} \quad \textbf{..... Eq. 7.10}$$

Question Q 7(3)

A bolus of 6 g drug is given IV and plasma concentration of the drug determined at intervals giving the following data:

Time since dose (h)	Plasma concentration (mg/L)
2.5	32
5	10

What is the half-life of the drug?

Time (*t*)	Concentration (*Cp*)	ln. *Cp*
0	100	4.61
1	50	3.91
2	25	3.22
3	12.5	2.53
4	6.25	1.83
5	3.125	1.14

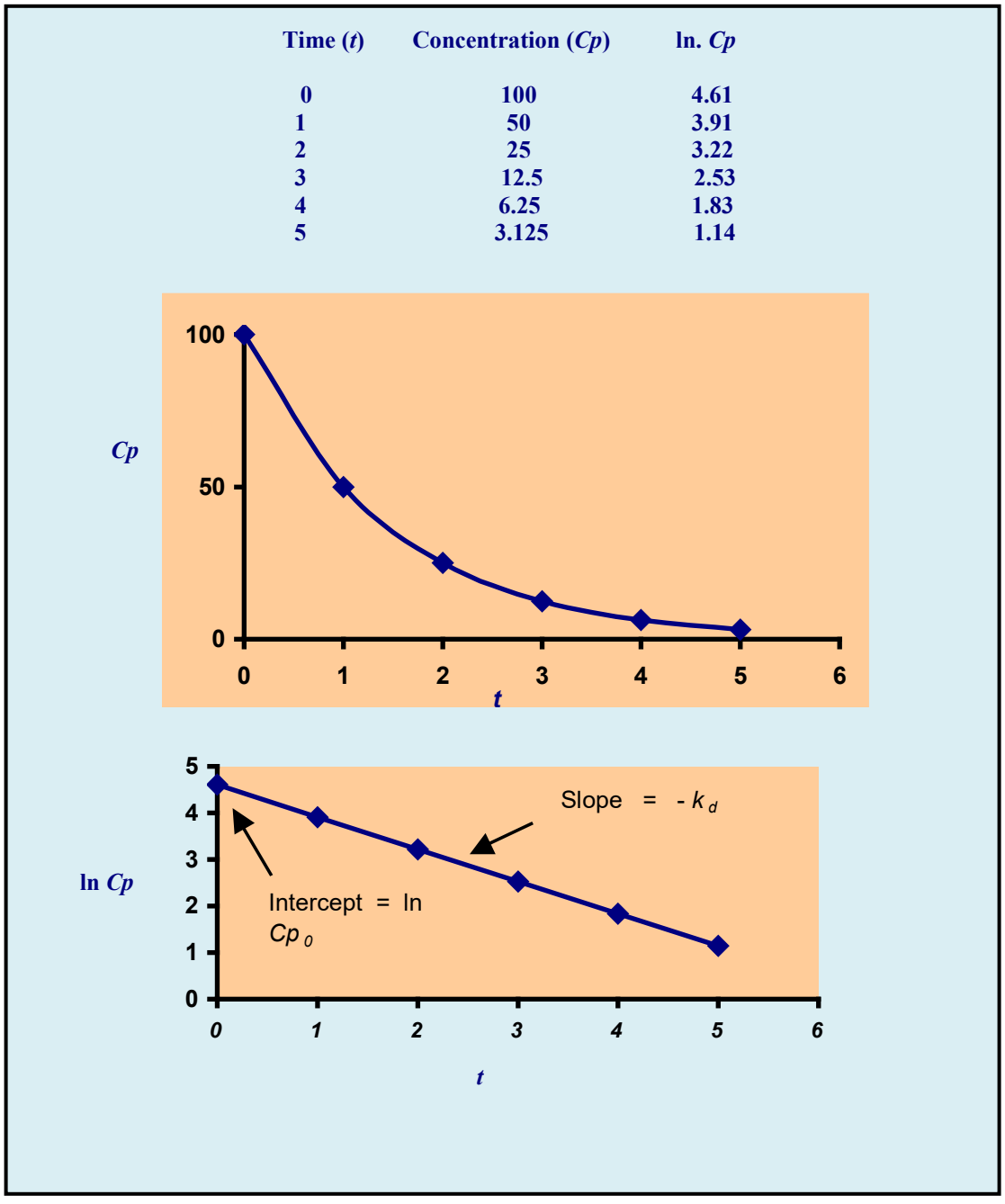


Figure 7.1 Clearance of a drug (concentrations expressed as percentage of initial value) with time (expressed as number of half-lives)

Answer Q 7(3)

The 2.5 h blood sample can be considered as the initial sample (Cp_0) and the 5 h sample as the sample when t = 2.5 h (i.e. 5.0-2.5) with concentration Cp_t. Therefore:

$$\begin{aligned} Cp_0 &= 32 \text{ mg/L} \\ Cp_t &= 10 \text{ mg/L} \\ t &= 2.5 \text{ h} \end{aligned}$$

These values can then be substituted into Eq. 7.10 and solved for $t_{½}$:

$$t_{½} = \frac{0.693.\ t}{\ln Cp_t - \ln Cp_0}$$

$$t_{½} = \frac{0.693 \times 2.5}{\ln 32 - \ln 10}$$

$$t_{½} = \frac{1.73}{3.47 - 2.30}$$

$$t_{½} = \frac{1.73}{1.17} = \textbf{1.5 h (2 sig figs)}$$

Note that ln *Cp* means the natural logarithm of concentration and not ln multiplied by *Cp*. Therefore ln Cp_t - ln Cp_0 is NOT the same as ln (Cp_t - Cp_0).

An alternative approach to this problem would be to plot the natural logarithm of the two drug concentrations (32 and 10 mg/L) against the times (2.5 and 5 h), measure the slope to derive k_d, then divide 0.693 by k_d to obtain the half-life.

So far we have been dealing with the kinetics of elimination following administration of a *single dose* of a drug. In many situations patients receive maintenance therapy: that is the drug is taken at regular intervals. The dose is repeated well before the previous dose has been eliminated from the body. Eventually, after multiple dosing, the situation is reached where the rate of administration of a drug (R_A) is equal to the rate of elimination (R_E) so that a constant *steady state* plasma concentration (Cp_{ss}) is achieved. The clearance of a drug (Cl) can be defined as the volume of plasma from which the drug is completely cleared per unit of time. Therefore the rate of elimination is the product of clearance and plasma steady state concentration:

$$R_E = Cl \times Cp_{ss}$$

The rate of administration is the amount of drug administered per unit time and depends on the dose administered, bioavailability (F), salt factor (S) and interval between doses (τ):

$$R_A = \frac{Dose \times F \times S}{\tau}$$

When a steady state is reached, $R_A = R_E$, and substitution of the expressions for R_A and R_E gives:

$$\frac{Dose \times F \times S}{\tau} = Cl \times Cp_{ss}$$

which can be rearranged to give an expression for the calculation of clearance:

$$\boldsymbol{Cl = \frac{Dose \times F \times S}{\tau \times Cp_{ss}}} \quad \textbf{..... Eq. 7.11}$$

For a drug which is eliminated by glomerular filtration alone, its clearance is equal to the patient's *GFR* (se chapter 5).

For drugs which exhibit first order kinetics, the elimination rate constant (k_d) can be defined as the fraction of the drug which is cleared in unit time:

$$k_d = \frac{\text{Amount of drug cleared in unit time}}{\text{Total amount of drug}} = \frac{R_E}{A_b}$$

Where A_b is the amount of drug in the body and is given by:

$$A_b = V_d \times Cp_{ss}$$

The rate of excretion (R_E) $= Cl \times CP_{ss}$

Substituting these values into the expression for k_d gives the following:

$$k_d = \frac{Cl \times Cp_{ss}}{V_d \times Cp_{ss}}$$

Canceling the Cp_{ss} terms gives a relationship between V_d, clearance and k_d:

$$k_d = \frac{Cl}{V_d} \quad \text{..... Eq. 7.12}$$

Practical applications of pharmacokinetics

The equations developed so far can be applied to answer many practical problems in drug therapy.

1. Calculation of plasma concentration any time after a loading dose or bolus of a drug is given.

First the theoretical initial plasma concentration (Cp_0) is calculated from the dose and volume of distribution (V_d) by substituting the expression for total amount absorbed (Eq. 7.4) into the rearranged expression for V_d (Eq. 7.5):

$$\text{Amount of drug in body} = Dose \times S \times F$$

$$\text{Plasma concentration } (Cp_0) = \frac{\text{Amount of drug in body}}{\text{Total volume } (V_d)}$$

$$Cp_0 = \frac{Dose \times S \times F}{V_d} \quad \text{..... Eq. 7.13}$$

The values for Cp_0, t and k_d (if k_d is not known it can be calculated from clearance and V_d using Eq. 7.12) can be substituted into Eq. 7.7 and solved for Cp_t.

Question Q 7(4)

A 70 kg lady is given an oral dose of carbamazepine of 400 mg. What is the plasma carbamazepine concentration 24 h later assuming a volume of distribution of 1.0 L/kg, a clearance of 0.05 L/h/kg, salt factor of 1 and a bioavailability of 0.75.

Answer Q 7(4)

First calculate the theoretical initial drug concentration (Cp_0) by substituting values for *dose*, *S*, *F* and V_d into Eq. 7.13:

$$Cp_0 = \frac{Dose \times S \times F}{V_d}$$

Where: $Dose$ = 400 mg
S = 1
F = 0.75
V_d = 1.0 L/kg = 1.0 x 70 = 70 L for patient's weight

$$Cp_0 = \frac{400 \times 1 \times 0.75}{70} = 4.29 \text{ mg/L}$$

Next substitute values for Cp_0, t and k_d (which must first be calculated from V_d and the clearance) into Eq. 7.7 and solve for Cp_t:

$$\ln Cp_t = \ln Cp_0 - k_d . t$$

Where: Cp_t = drug concentration at 24 h post dose = ?
Cp_0 = theoretical initial drug concentration = 4.29 mg/L
t = time since dose = 24 h
k_d = elimination rate constant calculated using Eq. 7.14:

$$k_d = \frac{Cl}{V_d} = \frac{0.05}{1.0} = 0.05 \text{ h}^{-1}$$

NB since the ratio of Cl to V_d is being calculated it is not necessary to correct these values for body weight. The units must however be compatible.

$$\ln Cp_t = \ln 4.29 - (0.05 \times 24)$$

$$\ln Cp_t = 1.46 - 1.20 = 0.26$$

$$Cp_t = \text{antilg}_e\ 0.26 = \textbf{1.3 mg/L (2 sig figs)}$$

2. Calculation of maintenance dose

Sometimes it is desirable to know which dose to give a patient (to be repeated at a time interval, τ) to achieve a target steady state plasma concentration (Cp_s). This is easily calculated by rearranging Eq. 7.11 to give:

$$\textit{Maintenance dose} = \frac{CP_{ss} \times Cl \times \tau}{S \times F} \quad \ldots\ldots\ldots\ldots \text{Eq. 7.14}$$

It is important to appreciate that a new a steady state will not be achieved until at least 5 half lives have elapsed.

3. Effect of a change in maintenance dose

This is simply calculated by inserting the new dose into Eq. 7.14. Alternatively the old steady state concentration can be multiplied by the fractional increase in dose.

$$\text{New } Cp_{ss} = \text{Old } Cp_{ss} \times \text{fractional change in dose}$$

Alternatively to calculate the dose necessary to change the steady state concentration by a set amount all that is necessary is to calculate the maintenance dose (using Eq. 7.14) required to achieve the difference in steady state concentration between the original value and the target level then add this to the original maintenance dose. Again it is necessary to wait at least 5 half-lives before the new steady state is achieved.

Question Q 7(5)

A 60 kg patient requires phenobarbitone to be given at 12 hourly intervals. Calculate:

a) The dose required to achieve an average steady state plasma concentration of 25 mg/L

b) The new average steady state plasma phenobarbitone concentration if the dose was increased to 120 mg.

Assume a clearance of 5 mL/h/kg and that both bioavailability and salt conversion factors are 1.

Answer Q 7(5)

a) The maintenance dose is calculated by substituting values for Cp_{ss}, t and clearance into Eq. 7.14:

$$\text{Maintenance dose} = \frac{Cp_{ss} \times Cl \times \tau}{S \times F}$$

Where Cp_{ss} = target average steady state plasma concentration = 25 mg/L
τ = dosing interval = 12 h
S = salt conversion factor = 1
F = bioavailability = 1
Cl = clearance = 5 mL/h/kg
= 5 x 60 = 300 mL/h/60 kg
= $\frac{300}{1000}$ = 0.3 L/h/60 kg

$$\text{Maintenance dose} = \frac{25 \times 0.3 \times 12}{1 \times 1}$$

= **90 mg**

a) If the dose is increased from 90 mg to 120 mg, then the fractional increase is given by:

$$\frac{\text{New dose - old dose}}{\text{Old dose}} = \frac{120 - 90}{90} = 1.33$$

The new average steady state concentration can be calculated by multiplying the old average steady state plasma concentration by this fractional increase in dose:

New Cp_{ss} = Old Cp_{ss} x Fractional increase in dose

= 25 x 1.33

= **33 mg/L (2 sig figs)**

4. Calculation of a loading dose (*LD*)

Equation Eq. 7.5 can be re-arranged to give an expression for the amount of drug in the body:

Amount of drug in body = Volume of distribution (V_d) x Plasma concentration (Cp)

The amount of drug in the body is the dose reaching the circulation so that Eq 7.4 can be substituted for the left hand side of this equation:

$$F \times S \times \text{Dose administered} = V_d \times Cp$$

The dose administered is then the loading dose (LD) required to achieve the desired plasma concentration of the drug (C_p) which can be rearranged to give the following expression:

$$\textbf{Loading dose } (LD) = \frac{V_d \times Cp}{F \times S} \quad \text{..... Eq. 7.15}$$

If the patient is already receiving the drug in question then this formula is modified to take account of the pre-existing drug concentration. The loading dose is then the amount of drug which needs to be administered to raise the plasma concentration from the initial concentration ($Cp_{initial}$) to the desired concentration ($Cp_{desired}$):

$$\textbf{Loading dose } (LD) = \frac{V_d \times (Cp_{desired} - Cp_{initial})}{F \times S} \quad \text{..... Eq. 7.16}$$

Question Q7(6)

Calculate the loading dose of digoxin (bioavailability = 0.75, salt factor = 1) required to achieve an initial plasma concentration of 1.5 µg/L in a 60 kg man (assume volume of distribution = 7 L/kg):

a) If the patient has never taken digoxin
b) If the patient is currently on digoxin with a plasma concentration of 0.5 µg/L.

Answer Q 7(6)

First calculate the volume of distribution (V_d) for the patient:

$$V_d = V_d \text{ (L/kg)} \times \text{Body weight (kg)}$$

$$= 7 \times 60$$

$$= 420 \text{ L}$$

Formula used to calculate loading dose:

$$LD = \frac{V_d \times (Cp_{desired} - Cp_{initial})}{F \times S}$$

a) If patient is not already on digoxin then $Cp_{initial}$ is zero:

$$LD = \frac{420 \times 1.5}{0.75 \times 1} = \mathbf{840\ \mu g}$$

b) If the patient is already on digoxin and $Cp_{initial}$ is 0.5 µg/L:

$$LD = \frac{420 \times (1.5 - 0.5)}{0.75 \times 1} = \mathbf{560\ \mu g}$$

5. Calculation of time to reach a steady state

Figure 7.1 shows that following a single dose of a drug its concentration falls to 50% of the original value after one half-life, 25% after two halve lives, 12.5% after three half lives etc. After five half lives the amount remaining is negligible at 3.25%.

Consider the situation when the same dose of a drug is administered using a dosing interval of one half-life. If the theoretical maximum plasma concentration achieved (Cp_0) is 100%, then after one half-life the concentration of the original dose will be 50%. Administration of a second dose at the end of one half-life contributes 50% to the plasma

concentration when two half lives have passed since the first dose, whereas the contribution at this time from the first dose will be one half of 50% which is 25%. Therefore after two halve lives the plasma concentration will be 50% + 25% = 75% of the theoretical maximum (i.e 75% of the steady state concentration). If this process is repeated for 5 half lives then the following pattern emerges:

	Number of half lives since initial dose				
Dose	**1**	**2**	**3**	**4**	**5**
1	50	25	12.5	6.25	3.125
2		50	25	12.5	6.25
3			50	25	12.5
4				50	25
5					50
Total	**50**	**75**	**87.5**	**93.8**	**96.9**

Figure 7.2 Cumulative drug concentration (expressed as percentage of theoretical maximum) for a drug administered at a dosage interval equal to its half-life

The concentration when a steady state is reached is given by the expression:

$$Cp_{ss} = \frac{Cp_0}{2} + \frac{Cp_0}{4} + \frac{Cp_0}{8} + \frac{Cp_0}{16} + \frac{Cp_0}{32} + \frac{Cp_0}{64} + \ldots\ldots$$

For practical purposes the plasma concentration after five half lives have elapsed approximates to the average steady state concentration. A similar argument, although more complex can be applied if the dosage interval is shorter than the half-life. A dosage interval of less than the half-life minimises oscillations around the average steady state concentration. If the dosage interval is considerably longer than the half-life then a steady state is never achieved and virtually all of the drug is removed from the body between each dose.

6. Calculation of pharmacokinetic parameters from a patient's data

Sometimes reliable pharmacokinetic data is not available or there may be considerable patient to patient variation in the handling of a drug. In this situation it is often possible to estimate kinetic parameters using the patient's own data. Sometimes a single test dose is employed to do this. It is vital however that exact details of the dose and timings of blood samples is accurately recorded. Provided clearance of the drug follows first order kinetics then it is possible to derive all kinetic parameters of interest from measurements made on two blood samples drawn at appropriate times.

Suppose that the dose is given at time t_0 and that samples are drawn at times t_1 and t_2 and that the measured plasma concentrations of these latter two samples are Cp_1 and Cp_2 respectively. The elimination rate constant can then be calculated by substituting these values into Eq. 7.8:

$$k_d = \frac{(\ln Cp_1 - \ln Cp_2)}{t_2 - t_1}$$

The half-life can then be calculated by substituting the value for k_d into Eq. 7.9:

$$t_{½} = \frac{0.693}{k_d}$$

The theoretical initial plasma concentration (Cp_0) can then be calculated by substituting either pair of concentration and times (it doesn't matter which) into the logarithmic integrated first-order rate equation (Eq. 7.7), together with the value for k_d then solving for Cp_0:

$$\ln Cp_t = \ln Cp_0 - k_d . t$$

$$Cp_0 = \text{antilog}_e (\ln Cp_0 - k_d . t)$$

The volume of distribution (V_d) is obtained by substituting the value of Cp_0 and the *dose* into Eq. 7.13:

$$Cp_0 = \frac{Dose \times F \times S}{V_d}$$

Which can be rearranged to give a value for $V_d / (F \times S)$:

$$\frac{V_d}{F \times S} = \frac{Dose}{Cp_0}$$

Note that F and S are unknown but are incorporated into the apparent value for V_d. Dosage adjustments can be made using the apparent V_d without knowledge of the actual values for F and S.

Substitution of k_d and $V_d/ (F \times S)$ into Eq. 7.12 allows calculation of clearance (or more precisely $Cl / (F \times S)$):

$$\frac{Cl}{F \times S} = k_d \times \frac{V_d}{F \times S}$$

The *maintenance dose* at dosage interval τ to achieve an average steady state drug concentration Cp_{ss} can then be obtained from Eq. 7.14.

$$\textit{Maintenance dose} = \frac{Cp_{ss} \times Cl \times \tau}{F \times S}$$

Again values for F and S do not need to be determined since they are already incorporated into the apparent clearance $Cl / (F \times S)$.

Question Q 7(7)

A 60 kg patient is given 4 g of a drug and blood samples drawn at timed intervals with the following results:

Time (h)	Plasma drug concentration (mg/L)
2	30
3	7

a) What is the half-life of the drug?
b) Calculate the apparent volume of distribution in L/kg.
c) Calculate the apparent clearance in L/h/kg.
d) What maintenance dose (to be administered once daily) would be needed to achieve an average plasma steady state concentration of 15 mg/L?

Answer Q 7(7)

a) To calculate the half-life first calculate the elimination rate constant using the logarithmic form of the integrated first-order rate equation (Eq. 7.7):

$$\ln Cp_t = \ln Cp_0 - k_d . t$$

Use the 1st (2 h) sample as the initial sample so that Cp_0 = 30 mg/L
Use the 2nd (4 h) sample as Cp_t (7 mg/L)
Therefore t = 4 - 2 = 2 h

$$\ln 7 = \ln 30 - k_d . 2$$

$$1.95 = 3.40 - 2\,k_d$$

$$2\,k_d = 3.40 - 1.95 = 1.45$$

$$k_d = \frac{1.45}{2} = 0.725\ h^{-1}$$

The half-life ($t_{1/2}$) is then calculated from k_d using Eq. 7.9:

$$t_{1/2} = \frac{0.693}{k_d} = \frac{0.693}{0.725} = \mathbf{0.96\ h}$$

b) First calculate the theoretical value for Cp_0 by substituting values for k_d, Cp_t and t (either set of t and Cp_t values can be used) into Eq. 7.7 but this time use the actual sample times.

Using the 2h results, t = 2h, Cp_t = 30 mg/L

$$\ln 30 = \ln Cp_0 - 2\,k_d$$

$$3.40 = \ln Cp_0 - 2 \times 0.725$$

$$\ln Cp_0 = 3.40 + 1.45 = 4.85$$

$$Cp_0 = \text{antilog}_e\ 4.85 = 128\ \text{mg/L}$$

Then calculate the apparent volume of distribution using Eq. 7.13.

$$Cp_0 = \frac{Dose \times F \times S}{V_d}$$

Which can be rearranged to give a value for the apparent volume of distribution:

Apparent volume of distribution i.e $\frac{V_d}{F \times S} = \frac{Dose}{Cp_0}$

Substitute Cp_0 = 7.03 mg/L and dose = 4 g = 4000 mg

$$\frac{V_d}{F \times S} = \frac{4000}{128} = 31.3 \text{ L}$$

Divide by body weight (60 kg) to give V_d in trerms of L/kg:

$$\frac{V_d}{F \times S} = \frac{31.3}{60} = \mathbf{0.52 \text{ L/kg}}$$

c) The apparent clearance can be calculated by substituting k_d and V_d into Eq. 7.12:

Apparent clearance i.e. $\frac{Cl}{F \times S} = k_d \times \frac{V_d}{F \times S}$

$$= 0.725 \times 0.52$$

$$= \mathbf{0.38 \text{ L/h/kg}}$$

d) The maintenance dose required to achieve an average steady state plasma concentration (Cp_{ss}) of 15 mg/L can be obtained by substituting values into Eq. 7.14:

$$\textit{Maintenance dose} = \frac{Cp_{ss} \times Cl \times \tau}{F \times S}$$

Cp_{ss} = 15 mg/L

$Cl / (F \times S)$ = 0.38 L/h/kg = 0.38 x 60 = 23 L/h/kg

τ = dosing interval = 24 h

Maintenance dose = 15 x 23 x 1 = **345 mg**

7. More complex situations

Although many drugs follow the principle outlined above (i.e. single compartment models obeying first-order kinetics) there are some notable exceptions. These drugs are often cleared by saturable mechanisms so that a concentration is reached at which the rate of elimination becomes independent of concentration. This is very important clinically since only small increases in dose above a threshold value can easily lead to toxic levels of the drug. A good example of this is phenytoin the clearance of which displays a mixture of zero and first-order kinetics. The best model to use under these circumstances is the equation of Michaelis and Menton which describes the rate of an enzyme reaction (v) in terms of substrate concentration (s) in terms of two constants the K_m (which can be shown to be the substrate concentration at half-maximal velocity) and the maximal velocity (V_{max}):

$$v = \frac{V_{max} \times s}{K_m + s}$$

Substituting the administration rate (which in a steady state is equal to the rate of elimination) for v:

$$v = \frac{Dose \times S \times F}{\tau}$$

and the average steady state drug concentration (Cp_{ss}) for s, gives the expression:

$$\frac{Dose \times S \times F}{\tau} = \frac{V_{max} \times Cp_{ss}}{K_m + Cp_{ss}}$$

which can be rearranged to give an equation to calculate the dose required to give the desired average steady state plasma drug concentration for the dosing interval:

$$\boldsymbol{Dose = \frac{V_{Max} \times Cp_{ss} \times \tau}{(K_m + Cp_{ss}) \times S \times F}} \quad \text{..... Eq. 7.17}$$

ADDITIONAL QUESTIONS

1. An antidepressant drug has a biological half-life of 30 hours. How long will it take a plasma concentration of 50 mg/L to fall to 20 mg/L?

2. A 15 year old boy presents to casualty following a convulsion. It turns out that he had swallowed 30 of his mother's lithium tablets about 10 hours previously. On admission his lithium concentration is 4.1 mmol/L. A decision needs to be made whether to haemodialyse him to reduce the lithium concentration. As this is not going to be available quickly, the physicians want to know how long he will have toxic levels just with endogenous clearance. Estimate the following, indicating clearly any assumptions you have made:

   a) The likely volume of distribution of the lithium at this stage in the situation, given a body weight of 65 kg.

   b) How long it will be before his lithium concentration drops to the relatively safe level of 1.5 mmol/L below which toxicity is unlikely, given a clearance of 0.03 L/h/kg.

3. A 60 mg dose of a drug is given to a male experimental subject who weighs 80 kg. Assuming that the drug is completely absorbed and distributed evenly throughout the total body water, estimate the potential peak plasma level. If the drug were distributed only within the extracellular compartment, what would the plasma level be?

4. A bolus of 6 g drug is given IV and 3 blood samples collected at intervals.

Time	mg/L
2.5h	32
5h	10
7.5h	3

   a) What is the half-life of the drug?
   b) What is the volume of its distribution?

5. The plasma concentration of a drug is found to be 200 nmol/L at 9.00 am. It's elimination follows first order kinetics with a rate constant is 0.34/h. Calculate the times at which the plasma concentrations will reach 100 nmol/L and 75 nmol/L.

6. A patient in casualty with a suspected adrenal crisis is given an IV dose of hydrocortisone at 18.00. The medical team on take wish to carry out a short synacthen test to confirm the diagnosis but there will be a significant contribution form the administered drug until its concentration has fallen to 10% of the peak value. If the half-life of hydrocortisone is 2 h, what is the earliest time at which the test can be carried out?

7. The SHO decides to treat a patient (weight 80 kg) with intravenous theophylline (salt factor = 0.8). What loading dose would you recommend in order to achieve a theophylline level of 12 mg/L given a volume of distribution of 0.5 L/kg and an initial plasma theophylline level of 4 mg/L?

8. A patient, unable to take oral medication, had been receiving intravenous valproate for a number of days and an average steady state level of 75 mg/L. After regaining consciousness the doctors wished two change to an oral twice daily regimen. In order to maintain the same average steady state concentration what dose would you recommend. Assume a clearance of 10 mL/h/kg, a bioavailability of 0.7 and a salt factor of 0.85.

Chapter 8

Body fluids and electrolytes

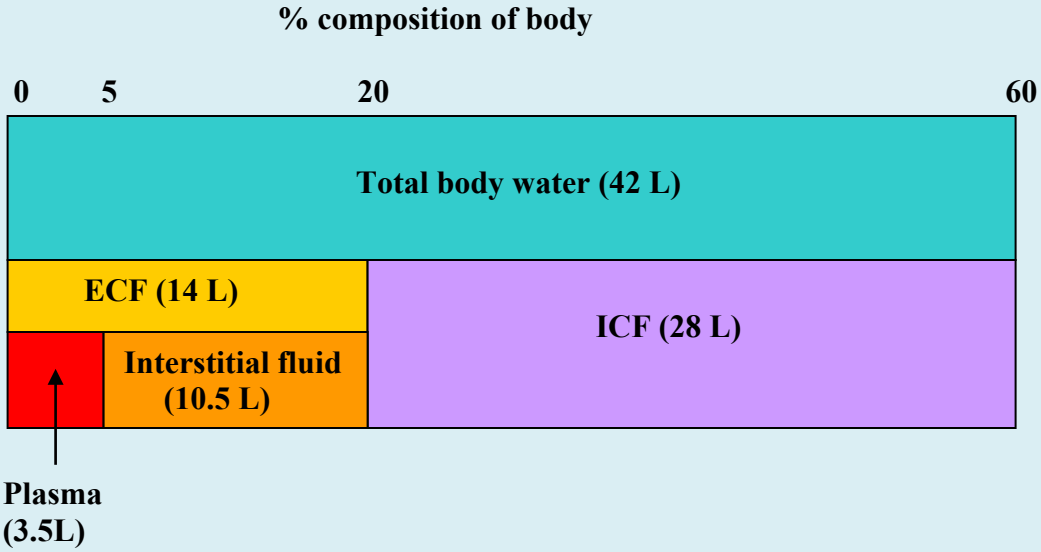
Distribution of body water

The human body contains approximately sixty per cent of its weight of water. For an average adult male weighing 70 kg this amounts to 42 L. Of this about a third (14 L) is contained in the *extracellular* fluid (ECF) and two thirds (28 L) in the *intracellular* fluid (ICF). The ECF and ICF are separated from each other by cell membranes which are semi-permeable but allow ready diffusion of water into and out of cells. Both of these compartments are in osmotic equilibrium with each other. Blood *plasma* constitutes approximately one quarter of the ECF (3.5 L), most of the remainder is present as *interstitial fluid* (10.5 L) which surrounds cells in the various tissues of the body. Plasma and interstitial fluid are separated by capillary walls which are freely permeable to water and electrolytes but minimally permeable to proteins such as albumin. The distribution of ECF between plasma and the interstitial fluid is controlled mainly by haemodynamic factors, integrity of capillary walls and the plasma albumin concentration. A small proportion of the ECF is present in the various body cavities (e.g. peritoneal, pleural, pericardial, synovial, spinal cavities). These relationships are illustrated in Fig 8.1.

Composition of body fluids

The fluid compartments differ in their solute composition. Sodium is the main cation in the ECF whereas potassium is the main cation within the ICF (see Fig 8.1). This status is maintained by the action of the sodium pump. The cell membrane of most tissues is impermeable to glucose in the absence of insulin. Since the interstitial fluid is created by filtration of plasma it is an ultrafiltrate of plasma with essentially the same solute composition but a much lower protein content. For practical purposes the concentration of glucose and electrolytes in the plasma and interstitial fluid can be regarded as identical.

a) Major fluid compartments of the body



b) Typical electrolyte compositions of the ECF and ICF

Electrolyte	Typical concentration (mmol/L)	
	ECF	ICF
Sodium	140	12
Potassium	4	145
Chloride	100	4
Bicarbonate	26	8

Figure 8.1 a) Distribution of body water between fluid compartments (ECF = extracellular fluid, ICF = intracellular fluid) – values shown for a typical adult male.
b) Typical electrolyte concentrations in the ECF and ICF.

Estimation of fluid losses

This can be extremely difficult since the mechanisms and consequences are complex and vary according to the nature of the fluid lost. In general however, several approaches can be used:

- Clinical assessment – such as presence of thirst, dry mouth, decreased sweating, tachycardia, decreased skin turgor, decreased urine output
- Loss of body weight – *short term* changes in body weight are due to changes in hydration. Difficult to assess in immobilised patients e.g. in ITU
- Fluid balance charts recording fluid input versus output – negative fluid balance consistent with fluid losses (this is not helpful if dehydration is already established). Accurate fluid balance charts are difficult to maintain and rely on estimating fluid loss via lungs, skin etc
- Laboratory measurements – haemoconcentration (raised haematocrit and serum proteins), increased serum creatinine and urea (with the increase in urea being disproportionate to creatinine), hypernatraemia, low volume concentrated urine (unless losses are due to renal losses).

A full discussion is beyond the scope of this book, but here are two areas where calculations are involved:

1. Calculation of fluid balance

When a patient is normally hydrated:

Net fluid intake = Net fluid losses

The difficulty is identifying and accurately quantifying all the fluid losses and gains. Small errors in the calculation of daily balance can accumulate over time with catastrophic consequences. If the net fluid loss exceeds the net fluid gain then the patient becomes water depleted i.e. is in negative fluid balance:

Fluid balance = Net fluid intake - Net fluid loss …….. Eq. 8.1

Fluid is gained by drinking. As food has a water content fluid is also gained by feeding, either orally, by naso-gastric tube feeding or by intravenous infusion. Another source of water gain which must be considered is water produced metabolically when fats and carbohydrates are oxidised. Fluid can be lost in urine, as water vapour via the lungs, sweating and evaporation through the skin and in faeces. In the hospitalised patient losses via the lungs may be increased if the patient is on a ventilator, losses via the skin may increased in burns patients or if the patient is pyrexial, and further losses may occur by vomiting, nasogastric suction and through drains and fistulae. Typical values for a normal adult in fluid balance are shown in figure 8.2.

Fluid gains		**Fluid losses**	
Oral (water)	**1200 mL**	**Urine**	**1600 mL**
Oral (food)	**800 mL**	**Lungs**	**300 mL**
Metabolism	**500 mL**	**Skin**	**400 mL**
		Faeces	**200 mL**
Total	**2500 mL**	**Total**	**2500 mL**

Figure 8.2. Typical gains and losses in a normal adult in perfect fluid balance

In practice it is difficult to measure the so-called *insensible losses* i.e. losses via the lungs, skin and faeces, and it is customary to assume a value of approximately 900 mL per day. Water production via metabolism, *the insensible gain*, is also difficult to determine, so an average value of about 500 mL per day is assumed. Therefore, *the net insensible loss* for an adult is normally in the order of 900 - 500 = 400 mL per day.

Question Q 8(1)

During a 24 h period a patient recovering from intestinal surgery receives 1.5 L of fluid by intravenous infusion. The total urine output during this period is 1200 mL and a further 450 mL of fluid was removed by nasogastric suction. Estimate the patients fluid balance.

Answer Q8(1)

Fluid gained			**Fluid lost**		
iv fluids	=	1500 mL	Urine	=	1200 mL
			Nasogastric	=	450 mL
			Insensible losses	=	400 mL
Total	=	1500 mL	Total	=	2050 mL

Fluid balance = Fluid gained - Fluid lost

= 1500 - 2050

= **- 550 mL**

Therefore the patient is a **negative** fluid balance of about **550 mL/24 h**.

2. Calculation of water loss from plasma sodium

The consequences of fluid loss depend upon the nature of the fluid lost. If the loss is isotonic, i.e. water and sodium are lost in the same proportion, due for example to haemorrhage, then the osmolality of the plasma (and ECF) remains unchanged and there is no stimulus for the osmotic shift of fluid from the ICF to the ECF. In other word acute loss of isotonic fluid is confined to the ECF. On the other hand if pure water loss occurs from the plasma (and ECF) then the osmolality (and hence sodium concentration) of the plasma (and ECF) rises and provides an osmotic stimulus for the shift of water from the ICF to the ECF until a point is reached when the two compartments are again in osmotic equilibrium. Note that at equilibrium the osmolality in both compartments will be higher than normal, but not as high as it would be if there had been no fluid shift from the ICF to ECF. In other word the loss in fluid volume is shared between the two compartments and so helps protect the circulating plasma volume (see Fig 8.3). In reality pure water loss rarely occurs, so calculation of the approximate fluid loss from changes in plasma sodium often serve as a (rough) *guide* for fluid replacement. The small difference between *osmolality* and *osmolarity* (see Chapter 6) will be ignored and it will be assumed that each mmol/L of solute contributes an osmolality of 1 mOsm/kg.

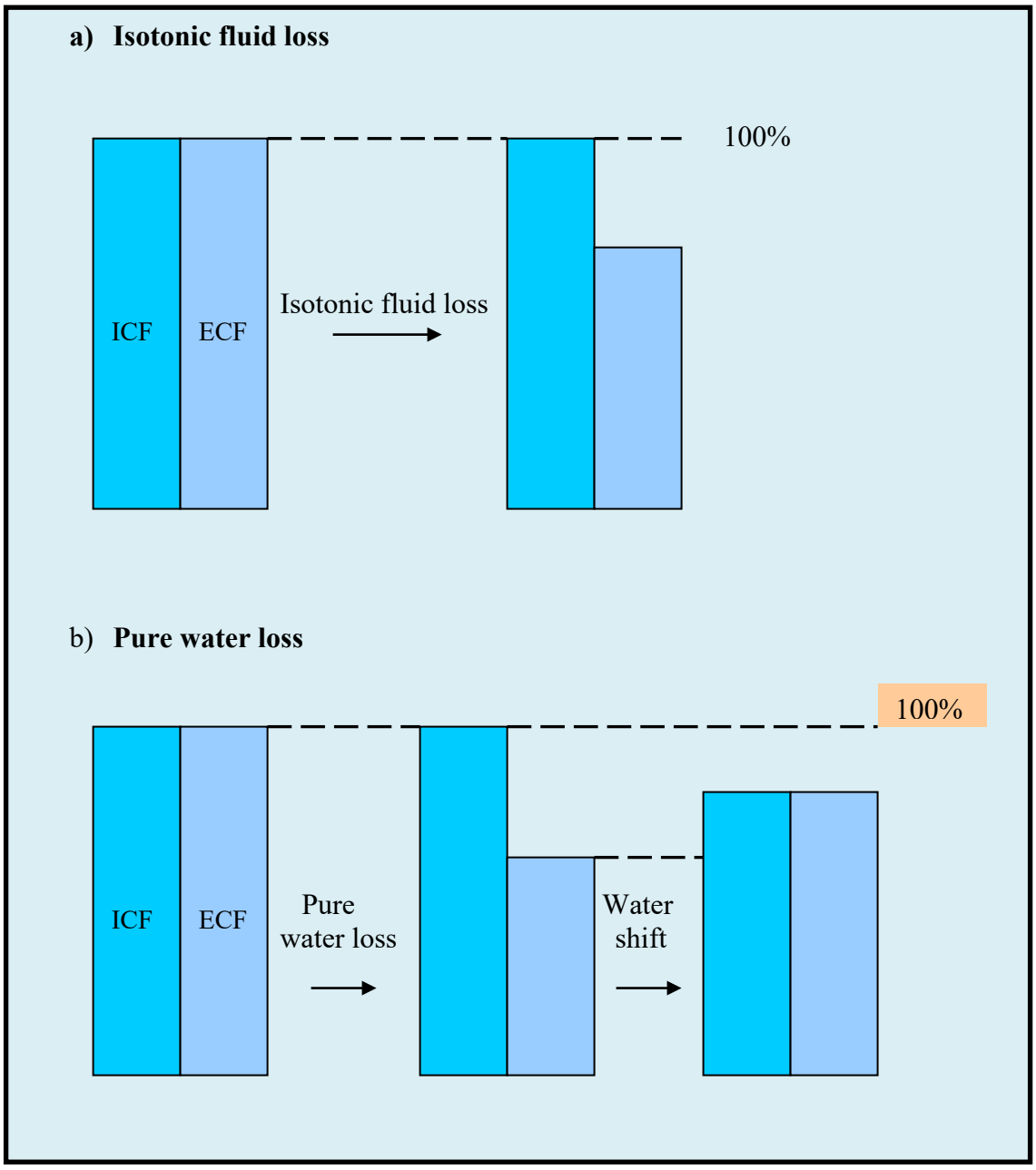


Figure 8.3 Comparison of the effects on ECF and ICF volumes of a) isotonic fluid loss, and b) pure water loss

If only pure water loss has occurred then the total amount of osmotically active species present in the body (both in the ICF and ECF) remains constant. The total amount of osmotically active species is the product of the volume of total body water and the plasma osmolality (the osmolality of all compartments must be the same since they are in osmotic equilibrium). Before the fluid loss occurred:

Total solutes (mOsm) = Initial osmolality (mOsm/kg) x Initial vol body water (L)

After loss of fluid and osmotic equilibrium has been reached:

Total solutes (mOsm) = Final osmolality (mOsm/kg) x Final vol body water (L)

Assuming no solutes have been lost then these two quantities are equal and we can write:

Final osmolality (mOS/kg) x Final vol (L)

= Initial osmolality (mOsm/kg) x Initial vol (L)

which can be rearranged to give an expression for the final body water volume:

$$\text{Final vol (L)} = \frac{\text{Initial osmolality (mOsm/kg)} \times \text{Initial vol (L)}}{\text{Final osmolality (mOsm/kg)}}$$

The volume of fluid lost is the difference between the initial volume and the final volume:

Fluid loss (L) = Initial volume (L) - Final volume (L)

Substituting for final vol gives:

Fluid loss (L) =

$$\textbf{Initial vol (L)} - \left[\frac{\textbf{Initial osmolality (mOsm/kg)} \times \textbf{Initial vol (L)}}{\textbf{Final osmolality (mOsm/kg)}}\right] \text{ ..... Eq. 8.2}$$

This formula can be simplified by taking the value for initial body water of 42 L (based on the average 70 kg male) and an average normal initial osmolality of 285 mOsm/kg:

$$\text{Fluid loss (L)} = 42 - \left\{\frac{285 \times 42}{\text{Final osmolality (mOsm/kg)}}\right\}$$

Which becomes:

$$\text{Fluid loss (L)} = 42 - \left\{ \frac{12000}{\text{Osmolality (mOsm/kg)}} \right\} \quad \text{..... Eq. 8.3}$$

Since sodium and its associated anions normally account for most of the osmolality of plasma then this expression can be further simplified by using the plasma sodium concentration assuming that it was initially normal (140 mmol/L), so that 140 x 42 = 5880 mmol:

$$\text{Fluid loss (L)} = 42 - \left\{ \frac{5880}{\text{Sodium (mmol/L)}} \right\} \quad \text{..... Eq. 8.4}$$

It cannot be emphasised too strongly that this formula can only give a crude estimate of fluid loss and is based on the following assumptions:

- That pure water loss has occurred
- That the plasma sodium was initially normal at 140 mmol/L (if the previous sodium concentration was known then it could be taken into account).
- There are no other abnormal amounts of osmotically active species present.
- There is no significant solute loss, gain or transfer between compartments.
- The initial body weight of the patient is 70 kg and contains 60 per cent water. If the body weight is known then a correction could be made, but this is more difficult for obese patients.

Question Q 8(2)

An adult dehydrated male has a plasma sodium of 165 mmol/L due to water depletion. Estimate the fluid deficit.

Answer Q 8(2)

Initial vol (L) x Intial Na (mmol/L) = Final vol (L) x Final Na (mmol/L)

Assume initial vol = 42 L; initial Na = 140 mmol/L
Final vol (L) = ?, Final Na = 165 mmol/L.

Substituting these values gives:

$$42 \times 140 = \text{Final vol (L)} \times 165$$

$$\text{Final vol (L)} = \frac{42 \times 140}{165} = 36 \text{ L (2 sig figs)}$$

Fluid loss (L) = Initial vol (L) - Final vol (L)

= 42 - 36 = **6 L**

2. Effect of hyperglycaemia on plasma sodium concentration

A rise in plasma glucose (for example due to insulin deficiency) is immediately accompanied by an increase in plasma (and hence ECF) osmolality which will stimulate the hypothalamic osmoreceptors which initiates an increase thirst and the release of antidiuretic hormone (ADH) from the posterior pituitary. What happens next depends upon whether the patient is able to drink an adequate amount of fluid (see Fig 8.4), but the end result is that there is a fall in plasma sodium concentration. It is often useful to predict this fall in sodium since if it does not account for the observed hyponatraemia then there must be another cause.

a) Assuming free access to fluid

The initial rise in osmolality due to the hyperglycaemia stimulates thirst via the hypothalamic osmoreceptors. The patient will respond to thirst by taking in fluids until a point is reached at which the plasma (and hence ECF) osmolality is returned to normal but at the expense of diluting sodium and other solutes i.e. to produce a dilutional hyponatraemia. Since the plasma osmolality is unchanged it follows that the rise in

plasma glucose concentration must be equal to the fall in sodium and its associated anions. Therefore the fall in sodium concentration must be equal to one half the rise in plasma glucose concentration:

$$\textbf{Fall in plasma sodium (mmol/L)} = \frac{\textbf{Rise in plasma glucose (mmol/L)}}{\textbf{2}} \quad \textbf{….. Eq. 8.5}$$

b) Assuming no intake of water

If there is no intake of water then the plasma osmolality will remain elevated (because the cell membranes are essentially impermeable to sodium and, when insulin deficient, impermeable to glucose). The osmotic stimulus will result in a shift of water from the ICF to the ECF (and hence plasma). An equilibrium is established in which the osmolality of both compartments is again equal but higher than normal. In other words the osmotic load is shared between both the ECF and ICF compartments. Therefore the increase in osmolality must be equal to the increase in plasma glucose concentration:

↑ plasma glucose = ↑plasma osmolality = ↑ECF osmolality = ↑ICF osmolality

The osmotic load in the ECF due to this accumulated glucose can be calculated from the rise in plasma glucose concentration (Δ glucose) and the ECF volume:

$$\text{Osmotic load (mOsm)} = \Delta \text{ glucose (mmol/L)} \times \text{ECF vol (L)}$$

At equilibrium the rise in osmolality (Δ osmolality) (which must be the same for both compartments) is given by the expression:

$$\Delta \text{ osmolality (mOsm/kg)} = \frac{\text{Osmotic load (mOsm)}}{\text{Total body fluid volume (L)}}$$

Substituting Δ glucose x ECF vol for the osmotic load and (ECF + ICF) for the total body fluid volume gives:

$$\Delta \text{ osmolality (mOsm/kg)} = \frac{\Delta \text{ glucose (mmol/L)} \times \text{ECF vol (L)}}{[\text{ECF vol (L)} + \text{ICF vol (L)}]}$$

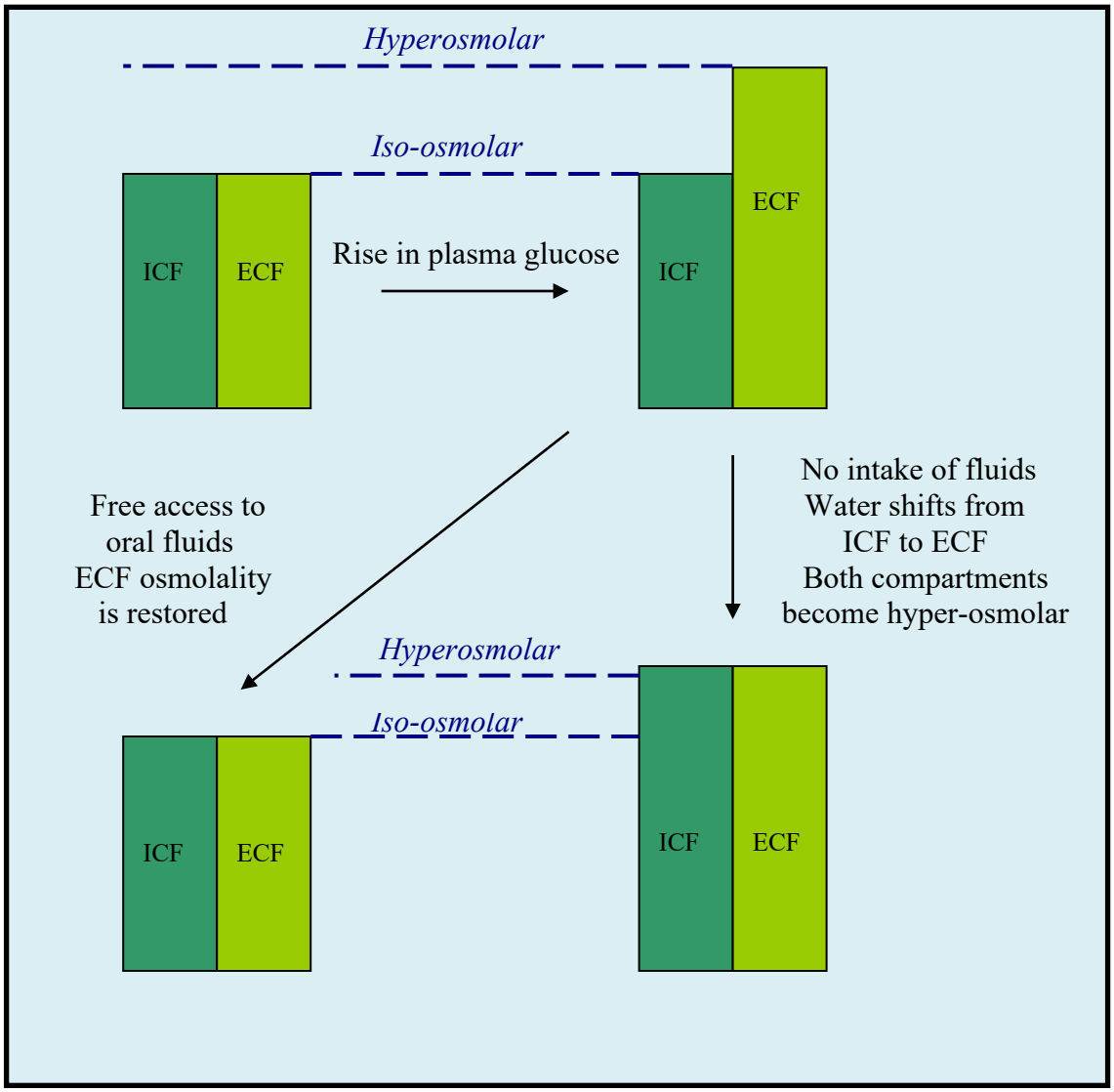


Figure 8.4 **Effect of a rise in plasma (and hence ECF) glucose on the osmolality of both the ECF and ECF compartments when there is either free intake of water or no intake at all. In both instances addition of water to the ECF compartment dilutes sodium and other electrolytes whilst the plasma glucose remains elevated**

Although some fluid has shifted between compartments, the ratio of (ECF + ICF) to ECF is still *approximately* 3:1 (since one third of body water is in the ECF). Therefore this expression becomes:

$$\textbf{Δ osmolality (mOsm/Kg)} = \frac{\textbf{Δ glucose (mmol/L)}}{\textbf{3}} \quad \textbf{..... Eq. 8.6}$$

Since the rise in osmolality is less than the rise in plasma glucose, the concentration of other solutes (predominantly sodium and its associated anions) must have fallen by an amount equal to the difference between the two:

$$\Delta (Na + Cl)\ (mmol/L) = \Delta \text{ osmolality } (mmol/L) - \Delta \text{ glucose } (mmol/L)$$

$$\Delta (Na + Cl)\ (mmol/l) = \{\Delta \text{ glucose } (mmol/L) / 3\} - \Delta \text{ glucose } (mmol/L)$$

Which is the same as:

$$\Delta (Na + Cl)\ (mmol/L) = - \frac{2 \times \Delta \text{ glucose } (mmol/L)}{3}$$

Since the concentration of sodium and its associated anions (mainly Cl^-) are assumed equal, then division by 2 gives the change in sodium concentration, which then cancels with 2 in the numerator, to give the expression:

$$\textbf{Δ Na (mmol/L)} = - \frac{\textbf{Δ glucose (mmol/L)}}{\textbf{3}} \quad \textbf{..... Eq. 8.6}$$

This calculation assumes that there has been no net gain or water loss by the body and no transfer of solutes between the ECF and ICF.

Question Q 8(3)

A male adult insulin dependent diabetic forgot to take his insulin. His blood glucose rose from 5 mmol/L to 20 mmol/L in 2 hours. During this period he did not pas any urine. Calculate the likely effect on his plasma sodium concentration assuming:

a) Free access to oral fluids
b) No intake of fluids occurs.

Answer Q 8(3)

a) Initially the increase in plasma glucose results in an increase in plasma osmolality. Hyperosmolality stimulates thirst and the patient drinks fluids which dilutes plasma until the stimulus is removed i.e. plasma osmolality has returned to normal. Since the plasma osmolality is unchanged but plasma glucose has risen it follows that the concentration of other solutes in plasma (principally sodium and its associated anions) must have fallen by an amount *equal* to the rise in glucose.

Fall in concentration of non-glucose solutes = rise in glucose concentration

= 20 - 5

= 15 mmol/L

Approximately one half of these solutes will be present as sodium, therefore:

$$\text{Fall in sodium} \quad \frac{15}{2} \quad = \quad \textbf{7.5 mmol/L}$$

b) Initially the increase in plasma glucose results in an increase in plasma osmolality. If there is no intake of fluids, then water moves by osmosis from the ICF to the ECF (which includes plasma) until equilibrium is established i.e. osmolalities of ICF and ECF are equal but elevated. The osmotic load is therefore shared between the two compartments.

Rise in amount of glucose in body (mmol)

= Rise (Δ) in plasma glucose concentration (mmol//L) x ECF vol (mmol/L)

This rise in the amount of glucose present in the body is responsible for the rise in overall osmolality. Therefore the increase in osmolality can be calculated by dividing this amount by the volume of total body water (i.e. ICF + ECF):

$$\Delta \text{ osmolality (mOsm/kg)} = \frac{\Delta \text{ glucose (mmol/L} \times \text{ECF Vol (L)}}{\text{ECF vol (L)} + \text{ICF vol (L)}}$$

Although there is small shift in fluid between compartments, approximately one third of body water is located in the ECF, the ratio ECF /(ECF + ICF) is roughly 3. Since the plasma glucose rose from 5 mmol/L to 20 mmol/L, the rise in osmolality is:

$$\Delta \text{ osmolality} = \frac{20 - 5}{3} = 5 \text{ mOsm/kg}$$

As the plasma osmolality has risen by 5 mOsm/kg and the plasma glucose has risen by 15 mmol/L, it follows that the concentrations of the other solutes in plasma (principally sodium and its associated anions) must have fallen by an amount equal to the difference between the two:

Fall in concentration of non-glucose solutes

$$= \Delta \text{ osmolality} - \Delta \text{ glucose}$$

$$= 5 - 15 = -10 \text{ mmol/L}$$

Approximately one half of these solutes will be present as sodium, therefore:

$$\text{Fall in sodium} \quad \frac{10}{2} = \mathbf{5 \text{ mmol/L}}$$

Direct versus indirect reading ion-specific electrodes

The advent of ion-specific electrodes for the determination of plasma/serum sodium soon led to discrepancies when compared with values determined by traditional flame photometry. The difficulty arose because these electrodes measure sodium *activity* in plasma *water*. Plasma contains appreciable protein (normally about 70 g/L). This protein, together with its hydration shell, occupies a significant proportion of plasma volume and reduces the amount of plasma water available to dissolve sodium and other ions. The "normal" plasma sodium determined by flame photometry is approximately 140 mmol/L of *plasma*. If this plasma contains 70 g/L of protein and we assume that its volume is approximately 70 mL/L (or 0.07 L/L), then this means that a litre of plasma contains only 1-0.07 = 0.93 L of water. Therefore the concentration of sodium in plasma water can be calculated as follows:

$$\text{Na (mmol/L water)} = \frac{140}{0.93} = 151 \text{ mmol/L}$$

This is the value which will be obtained for plasma sodium containing 140 mmol/L plasma when measured on whole undiluted plasma in an ion-specific electrode system. Since the measurement is made directly on whole plasma rather than diluted plasma it is said to be a "*direct reading* electrode". In general:

$$\mathbf{[Na^+]\ (plasma\ water) = \frac{[Na^+]\ (plasma)}{1 - Plasma\ protein\ (Kg/L)}} \quad \ldots\ldots\ldots \textbf{Eq. 8.7}$$

If, on the other hand, the plasma sample is first diluted with an aqueous diluent before the electrode measurement is taken then the discrepancy almost disappears. These instruments are known as "*indirect reading* electrodes" Consider the same sample, containing 140 mmol of sodium per L plasma, first diluted 1 in 20 before the sodium measurement is made. This is equivalent to diluting 0.05 L of plasma to 1 L. The sodium present in 1 L of plasma diluted 1 in 20 will be 140 x 0.05 = 7 mmol. The amount of water this sodium is dissolved in will be equal to the amount of water in the 0.05 L plasma sample plus the water added to it.

Water in 0.05 L plasma = 0.05 L - Volume due to protein

If the protein is 70 g/L (approx 0.07 L/L), then 0.05 L of plasma is occupied by 0.05 x 0.07 = 0.0035 L of protein. The plasma sample therefore contains 0.05 - 0.0035 = 0.0465 L of water. To dilute the plasma 1 in 20, the amount of diluent added is 1 - 0.05 = 0.95 L. Therefore:

Total amount of water in 1 L of 1 in 20 dilution of plasma = 0.95 + 0.0465 = 0.9965 L.

$$\text{Concentration of sodium} = \frac{7}{0.9965} = 7.025 \text{ mmol/L plasma water}$$

Allowing for the 1 in 20 dilution:

Plasma sodium = 7.025 x 20 = 140.5 mmol/L

The reason for this discrepancy is that dilution results in a decrease in the *proportion* of water displaced by protein. Plasma sodium measured with a flame photometer gives similar readings to indirect reading electrodes.

It has become common practice for instrument manufacturers to "adjust" the calibration of their direct reading ISE instruments so that the discrepancy with flame photometers disappears, but only at a "normal" plasma protein concentration (usually 70 g/L):

$[Na^+]$ (plasma) = $[Na^+]$ (water) x 0.93

If the plasma protein differs markedly from "normal" then the discrepancy with flame photometer readings reappears. In the above example, if the instrument is adjusted so that the plasma water sodium of 151 mmol/L reads 140 mmol/L then a sample with the same plasma sodium concentration but containing 50 g/L (0.05 kg/L) protein is measured, then the concentration of sodium in plasma water will be:

$$[Na^+] \text{ (plasma water)} = \frac{140}{1 - 0.05} = \frac{140}{0.95} = 147 \text{ mmol/L}$$

If the instrument makes the same adjustment (by assuming that the plasma protein is 70 g/L), then the reading given will be:

Reported $[Na^+]$ = 147 x 0.93 = 137 mmol/L.

Question Q 8(4)

A plasma contains 140 mmol/L of sodium and 95% water by volume. Neglecting sodium binding by plasma proteins, calculate the apparent plasma sodium concentration determined from measurements with an electrode system which responds to water sodium (a) in undiluted plasma, and (b) in plasma diluted 1 in 20 with water.

Answer Q 8(4)

a) Undiluted plasma – contains 95% water

Water concentration of plasma $= \frac{1 \times 95}{100} = 0.95$ L/L

Therefore concentration of sodium in plasma water is:

$[Na^+]$ (plasma water) $= \frac{140}{0.95} =$ **147 mmol/L**

b) Plasma diluted 1 in 20 with water:

Working on a volume of 1 L i.e. 50 mL (0.05 L) plasma diluted to 1 L with water:

Amount of sodium in 1 L diluted plasma $= \frac{140}{20} = 7$ mmol

Amount of water in diluted plasma = Water from plasma + water from diluent

$= \frac{(0.05 \times 95)}{100} + 0.95 = 0.9975$ L

$[Na^+]$ (plasma water) $= \frac{7}{0.9975} = 7.02$ mmol/L diluted plasma

Multiplication by the dilution factor of 20 gives the sodium concentration in undiluted plasma:

$[Na^+]$ in undiluted plasma $= 7.02 \times 20 =$ **140 mmol/L (3 sig figs)**

The anion gap

To maintain electrical neutrality the sum of concentrations of cations (sodium, potassium, calcium, magnesium etc) must equal that of all the anions (chloride, bicarbonate, phosphate, sulphate and proteins). Some, but not all of these are frequently measured in clinical practice. If only sodium chloride and bicarbonate are measured then the following relationship can be written:

$$[Na^+] + [\text{unmeasured cations}] = [Cl^-] + [HCO_3^-] + [\text{unmeasured anions}]$$

Which can be rearranged to:

$$[Na^+] = [Cl^-] + [HCO_3^-] + [\text{unmeasured anions}] - [\text{unmeasured cations}]$$

Since the concentrations of unmeasured anions always exceeds that of unmeasured cations, their difference is defined as the *anion gap*:

$$\text{Anion gap} = [\text{unmeasured anions}] - [\text{unmeasured cations}]$$

Substituting anion gap into the previous expression gives:

$$[Na^+] = [Cl^-] + [HC0_3^-] + \text{Anion gap}$$

which can rearranged to give the following expression for the anion gap:

$$\textbf{Anion gap} = [Na^+] - [Cl^-] - [HCO_3^-] \quad \text{..... Eq. 8.8}$$

The anion gap was originally developed as a quality control tool when it was noted that in most patients the difference between the sodium concentration and the sum of the chloride and bicarbonate concentrations was always approximately 12 mmol/L. Sometimes the potassium concentration is included in the calculation. Small deviations from the reference range for the anion gap (7-16 mmol/L) are usually due to marked changes in plasma calcium, potassium, phosphate and negatively charged proteins. However, the principal use of the anion gap is as an aid in the differential diagnosis of non-respiratory acidosis. A markedly raised anion gap indicates the presence of excess unmeasured anions of metabolic acids e.g. ketoacids, lactic, salicylic, oxalic (from metabolism of ethylene glycol) and formic acids (from metabolism of methanol).

Question Q 8(5)

The following results were obtained on a young adult in the Accident and Emergency Department:

Plasma sodium	=	140 mmol/L
Plasma chloride	=	97 mmol/L
Plasma bicarbonate	=	8 mmol/L

Calculate the anion gap.

Answer Q 8(5)

$$\text{Anion gap} = [Na^+] - \{[Cl^-] + [HCO_3^-]\}$$

$$= 140 - \{97 + 8\}$$

$$= 140 - 105$$

$$= \mathbf{35} \text{ mmol/L}$$

ADDITIONAL QUESTIONS

1. Over a 24 h period a patient recovering from intestinal resection receives 2 L of fluids intravenously and 750 mL orally but does not eat any solids over this period. The urine output over the same period is 1.25 L and 600 mL of fluid is lost via a fistula. Is the patient in positive or negative fluid balance and by how much?

2. A patient known to have diabetes insipidus is admitted in coma. His plasma osmolality is 324 mosm/kg. If his weight is 85 kg, estimate his body water deficit.

3. A male adult insulin dependent diabetic forgot to take his insulin. His blood glucose concentration, which was 5 mmol/L, rose to 15 mmol/|L in two hours. Estimate the effect on his plasma sodium concentration, assuming that no other water intake nor loss of water from the body takes place during this time, indicating what assumptions you make.

4. A plasma sample with a total protein content of 70 g/L gave identical sodium results of 140 mmol/L when measured using either a direct-reading ion-selective electrode or a flame photometer. What plasma sodium result would you expect the ion-selective electrode to give with the same plasma sample if its total protein concentration had been 90 g/L?

Chapter 9

Enzymology

Why measure enzyme activity?

Enzymes are proteins which catalyse chemical reactions. Enzymes are of interest to the clinical biochemist for a number of reasons:

- Enzymes are often released from tissues into the circulation as a result of disease e.g. the release of aspartate aminotransferase from the liver affected by hepatitis.
- Inherited diseases are often due to a deficiency in a particular enzyme e.g. glucose-6-phosphatase in type-I glycogen storage disease.
- Enzymes may be used as analytical tools e.g. hexokinase and glucose-6-phosphate dehydrogenase in the assay of glucose.

Several different approaches can be used to quantify enzymes:

- Measurement of the amount of enzyme protein following isolation from the analytical sample. This approach is seldom used since enzyme purification is a lengthy procedure during which some losses are inevitable.
- Since enzymes are proteins, immunoassay can be used (i.e. mass measurements). This approach has been used for the assay of the MB isoenzyme of creatine kinase.
- Measurement of the rate of the enzyme reaction (i.e. catalytic activity).

Catalytic activity

In routine clinical practice enzymes are usually quantified by measuring their catalytic activity. The rate of an enzyme reaction is dependent upon many factors (see Fig 9.1) but, with a few exceptions and provided the conditions of the assay are carefully chosen, the rate of the reaction is always proportional to the concentration of the enzyme in the reaction mixture. The rate (or activity) will vary depending upon the analytical conditions used and over the past few decades considerable effort has been expended by biochemists to standardise assay conditions so that activity measurements obtained in individual laboratories are comparable. Reaction conditions have been optimised so that the highest rate (maximal sensitivity) is obtained and small variations in conditions (substrate and cofactor concentrations, pH etc) have minimal effect. Whichever assay is used the rate is proportional to enzyme concentration but the actual rate depends on the reaction conditions employed.

- **Enzyme concentration**
- **Nature of the substrate and any cofactor(s)**
- **Concentration of substrate and any cofactors**
- **Buffer and its concentration**
- **pH**
- **Ionic strength**
- **Inhibitor/activator concentration**
- **Temperature**

Figure 9.1 Factors affecting enzyme activity

How is activity measured?

An enzyme catalyses the conversion of its substrate into a product. Therefore the course of the reaction can be followed by either measuring the disappearance of substrate or formation of product (Fig 9.2).

Rate of consumption of substrate(s) = Rate of formation of product(s)

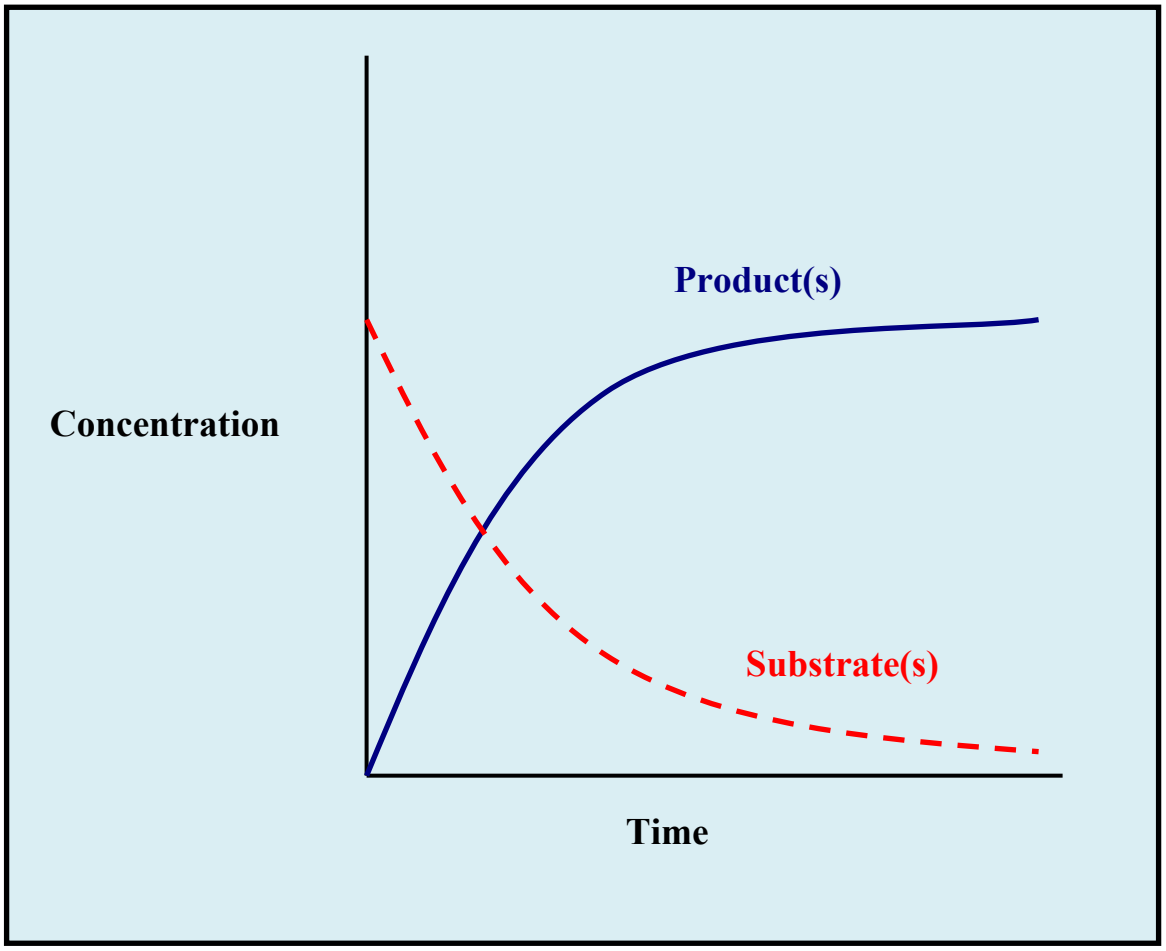


Figure 9.2 Progress curves of an enzyme catalysed reaction

If the product of the reaction does not have a physical property (such as absorbance) by which its appearance can be monitored then either a reagent is added to form a suitable derivative or a second enzyme is added to convert the product into another compound which is easily measurable. When a second (or third) enzyme is added in this way the assay is said to be *coupled*. It is vital that the concentration of the second enzyme is present *in excess* so that product is removed as soon as it is formed i.e. the first enzyme reaction (the one we wish to measure) is always rate limiting.

There are two approaches which can be used to obtain a rate measurement:

- Take measurements (of product or substrate concentration) at two points in time, calculate the difference then divide by the time period to give the rate (i.e. change of concentration per unit time). These are usually referred to as *fixed-time* methods. The term *end-point* assay is often used but is incorrect because enzyme activity is always a rate measurement and therefore needs at least two measurements. A second reading is always required and sometimes it is assumed (rightly or wrongly) that a reagent blank gives a reliable measure of concentration at time zero.

- Monitor the reaction continuously and take a rate measurement over a suitable time period. These *continuous monitoring* methods are preferred since it is possible to evaluate the reaction progress and ensure that a true initial rate measurement is taken and is constant, avoiding errors due to any lag-phase.

Whichever approach is used, the timing of the measurements is critical. Initially (or possibly after any lag-phase) the rate of reaction at any given enzyme concentration is constant. However, as the reaction progresses substrate is consumed and its concentration falls and eventually a point is reached at which substrate availability become rate limiting and the rate of the reaction falls i.e. the progress curve becomes non-linear. This is illustrated in Fig 9.3. Unless a lag-phase is observed, measurement over the segment *0-A* gives a true initial rate whereas measurement over the segments *0-B* or *0-C* gives an artificially low result. At point *C* the substrate is completely exhausted (or the reaction is at equilibrium) and the reading is actually a measure of substrate rather than enzyme concentration. Sometimes a lag phase is observed so that the initial rate is less than optimal (see inset to Fig 9.3). In this situation the optimal rate is given over the segment A-B not 0-A.

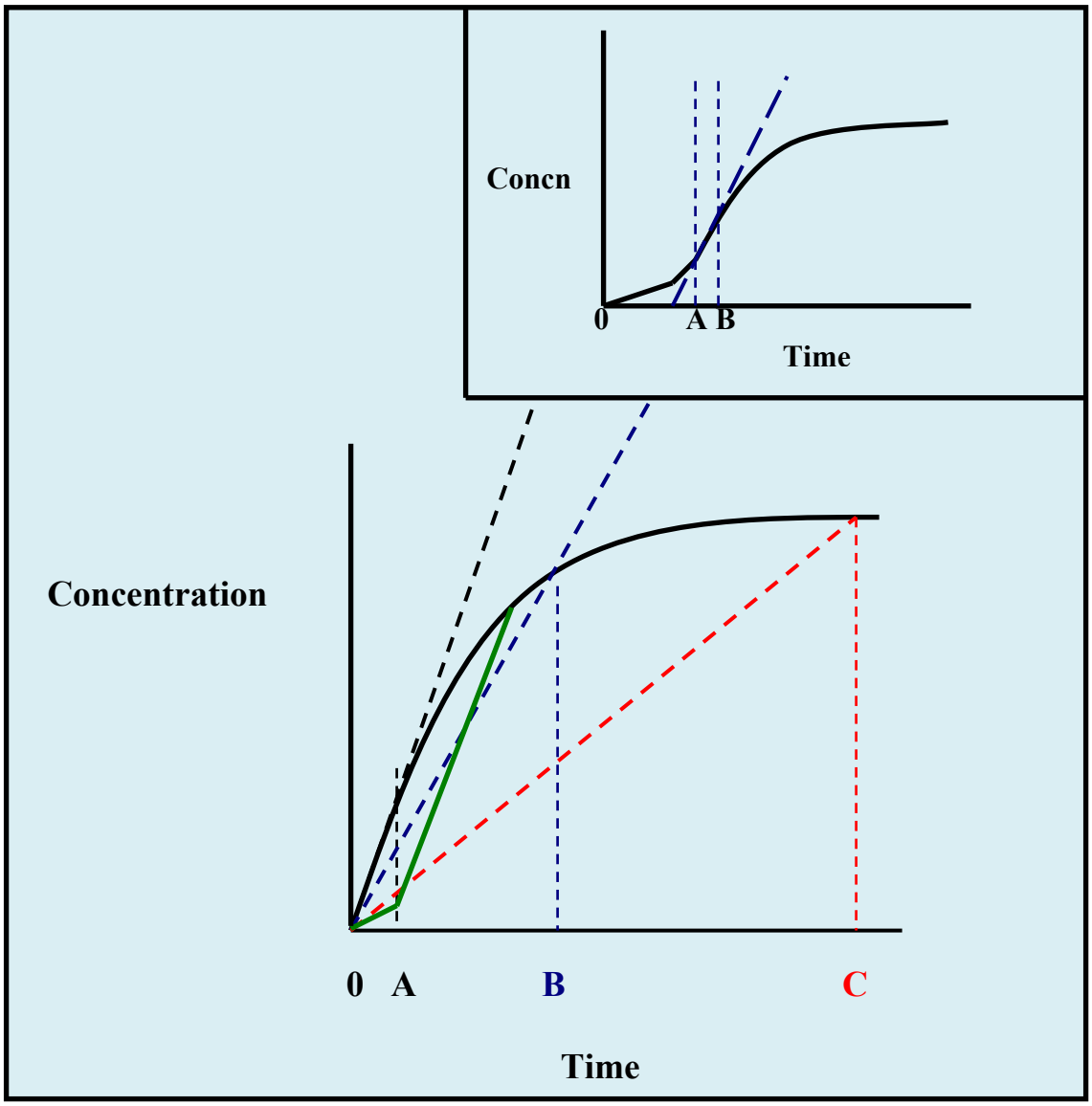


Figure 9.3 Effect of measurement period on rate of an enzyme-catalysed reaction. Inset shows an enzyme catalysed reaction with a lag-phase

Units for expressing enzyme activity

In order to make use of enzyme activity measurements for diagnostic purposes it is obviously essential to express the result in a way which makes comparison with a reference range or a patient's previous result easy. In the early days of diagnostic enzymology units were often named after the originator of the method used. For example

King-Armstrong (KA) units were used for alkaline phosphatase (the amount of enzyme present in 100 mL of serum that will split 1mg of phenol from phenylphosphate in 1 hour). Very soon a plethora of units were in use and attempts were made to standardise enzyme units. This may seem a rather pointless exercise since although units may be identical the enzyme activity will not be identical unless the reaction conditions are held constant. Nevertheless, the Commission on Enzymes of the International Union of Biochemistry propose that enzyme activity should be expressed in terms of *international units*. *One international unit (U) is the quantity of enzyme that catalyses the reaction of 1 µmol of substrate per minute.* Catalytic concentration is to be expressed in terms of U/L or mU/L, whichever gives the more convenient numerical value.

The international unit itself may eventually be replaced by a new unit termed the *katal* and concentrations expressed as katals per litre (kat/L). *One katal is the amount of enzyme which catalyses the reaction of 1 mol of substrate per second*.

When there is some uncertainty about the exact nature of the substrate (e.g. where the substrate is a macromolecule such as starch or a protein) then units are still expressed as the amount of a group or reside released per unit time (e.g. glucose units or amino acids formed per minute).

Calculating enzyme activity

This involves converting physical measurements (e.g. absorbances) made over timed interval(s) into substrate concentration units which are then used to derive a rate for the enzyme catalysed reaction. This rate then needs to be converted to the concentration of enzyme units in the clinical sample. The following example illustrates the process:

Question Q 9(1)

50 µL serum is added to 2 mL NADH solution (0.17 mmol/L) in Tris buffer (5.6 mmol/L) and incubated at 37°C for 10 min. 0.2 mL sodium pyruvate solution (13.5 mmol/L) is added and the rate of absorbance change monitored at 340 nm in a cuvette with a path-length of 0.5cm. The absorbance readings at 30 s and 60 s are 0.183 and 0.148 respectively. Calculate the LDH activity in the serum (molar absorption coefficient of NADH at 340 nm = 6.30×10^3 L/mol/cm).

Answer Q 9(1)

The reaction catalysed by lactate dehydrogenase (LDH) is:

$$\text{Pyruvate} + \text{NADH} + \text{H}^+ \rightarrow \text{Lactate} + \text{NAD}^+$$

This is the reverse of the normal reaction. The cofactor NADH absorbs at 340 nm, whereas absorbance due to NAD^+ is negligible. Therefore as the reaction progresses the absorbance at 340 nm falls at a rate which is equal to the rate of consumption of the substrate (pyruvate).

The concentration of NADH (c) at any point in time can be calculated from the absorbance reading (A), the cuvette path length (b) and the molar absorptivity of NADH (a) using equation Eq 4.4:

$$A = abc \qquad \text{rearranged to give} \qquad c = \frac{A}{ab}$$

Therefore at the first time the absorbance reading is taken (t_1 = 30 seconds), A = 0.183, b = 0.5 cm and a = 6.3 x 10^3 L/mol/cm so the concentration of NADH (c_1) in mol/L can be calculated as follows:

$$c_1 = \frac{0.183}{6.3 \times 10^3 \times 0.5} = 5.8 \times 10^{-5} \text{ mol/L}$$

The same calculation could be performed for the second absorbance reading (when t_2 = 60 seconds and A_2 = 0.148) to give the concentration (c_2) at time t_2. The two equations for the calculation of the concentrations at t_1 and t_2 are therefore:

At t_1: $\quad c_1 = \frac{A_1}{ab}$

At t_2: $\quad c_2 = \frac{A_2}{ab}$

These two expressions can be combined to calculate the change in concentration (Δc) over the time period $t_2 - t_1$ (Δt):

$$\Delta c = c_1 - c_2 = \frac{A_1}{ab} - \frac{A_2}{ab} = \frac{(A_1 - A_2)}{ab} = \frac{\Delta A}{ab}$$

(Mathematicians use the symbol Δ to denote a difference or change between two values.)

This gives the decrease in concentration of substrate over the time period (Δt) in units of mol/L. Since International Units use concentration expressed as µmol/L, this value must be multiplied by 1,000,000 (since there are 1,000,000 µmol in a mol):

$$\Delta c = \frac{\Delta A \times 1{,}000{,}000}{a \times b} \quad \mu\text{mol/L reaction mixture}$$

Division by the time interval ($t_2 - t_1 = \Delta t$) gives the rate of change in concentration as µmol/*sec*/L reaction mixture. Multiplication by 60 converts this rate to µmol/*min*/L reaction mixture:

$$\text{LDH activity} = \frac{\Delta c}{\Delta t} = \frac{\Delta A \times 1{,}000{,}000 \times 60}{\Delta t \times a \times b} \quad \mu\text{mol/min/L reaction mixture}$$

It is usual to express enzyme activity in serum as U/L of *serum*, not U/L of *reaction mixture.* Therefore the dilution of the serum in the reaction mixture needs to be taken into account. Multiplying by the total reaction volume and dividing by the sample volume gives:

$$\text{LDH activity} = \frac{\Delta A \times 1{,}000{,}000 \times 60 \times \text{Total reaction volume}}{\Delta t \times a \times b \times \text{Sample volume}} \quad \mu\text{mol/min/L serum}$$

Substituting:

$\Delta A = A_1 - A_2 = 0.183 - 0.148 = 0.035$
$\Delta t = t_2 - t_1 = 60 - 30 = 30$ seconds
a = molar absorptivity of NADH = 6.30×10^3 L/mol/cm
b = cuvette path length = 0.5 cm
Total reaction volume = 0.05(serum) + 2(NADH/buffer) + 0.2(substrate) = 2.25 mL
Sample volume (serum) = 0.05 mL

$$\text{LDH activity} = \frac{0.035 \times 1{,}000{,}000 \times 60 \times 2.25}{30 \times 6.30 \times 10^3 \times 0.5 \times 0.05} = \mathbf{10000\ U/L}$$

If a large number of calculations are to be performed for the same assay then it would be simpler to combine all the terms (except ΔA) to produce a factor which could then be used to obtain enzyme activity directly i.e. $\Delta A \times$ Factor (28600) = LDH activity (U/L).

1. Divide change in absorbance by the time period (in min) over which the measurements were taken to give rate i.e. ΔA/min
2. Divide by molar absorptivity (units: L/mol/cm) and cuvette path length (cm)
3. Multiply by 1,000,000 to convert from mol to μmol
4. Multiply by total reaction volume (mL) and divide by sample volume (mL)

For an enzyme assay utilizing NADH/NAD as cofactor (monitored at 340 nm) the formula used is:

$$\text{Enzyme activity (U/L)} = \frac{\Delta A/min \times 1{,}000{,}000 \times \text{Total vol (mL)}}{6.30 \times 10^3 \times \text{Path length (cm)} \times \text{Sample vol (mL)}}$$

which simplifies to:

$$\text{Enzyme activity (U/L)} = \frac{\Delta A/min \times 160 \times \text{Total vol (mL)}}{\text{Path length (cm)} \times \text{Sample vol (mL)}}$$

Figure 9.4 Steps in the calculation of enzyme activity

Conversion of enzyme units

The older literature is full of enzyme data expressed in units other than U/L. It is sometimes useful to convert these values to the corresponding activity in U/L. For example, King-Armstrong (KA) units were used for many years to report alkaline phosphatase activity. One KA unit is the amount of enzyme in 100 mL of serum that will split 1 mg of phenol from phenylphosphate in 1 hour and can be written:

1 KA unit = 1 mg phenol/h/100 mL serum

To convert to activity expressed as international units (U/L) it is necessary to determine the number of µmol of phenol formed in 1 min in 1 L of serum. The following steps are involved:

1. Multiply by 1,000 to convert mg to µg
2. Divide by the molecular weight of phenol (94) to convert from µg to µmol
3. Divide by 60 to convert reaction period from h to min
4. Multiply by 10 to convert from 100 mL serum to 1 L serum

The final result is: $1 \text{ KA unit} = \frac{1{,}000 \times 10}{94 \times 60} = 1.77 \text{ U/L}$

Therefore:

$$\text{Alk phos (U/L)} = \text{Alk phos (KA units)} \times 1.77$$

$$\text{Alk phos (KA units)} = \frac{\text{Alk phos (U/L)}}{1.77}$$

It is important to remember that even after converting enzyme activity from one unit to another, the numerical result will still depend on the reaction conditions used.

Question Q 9(2)

A transaminase result is quoted in the literature as 207 Karmen units. One Karmen unit is the amount of transaminase that will produce an absorbance change of 0.001 in a 1cm cuvette at 340 nm (a coupled reaction) in 1 min per 1 mL serum (in a total volume of 3 mL). Assuming the molar absorptivity of NADH is 6.30×10^3 L/mol/cm, express the transaminase activity as *a)* international units per L of serum, and *b)* katals/L.

Answer Q 9(2)

a) 1 Karmen Unit = 0.001 *A*/min/mL serum (in a total volume of 3 mL)

1 U/L = 1 µmol/min/L serum

Therefore:

1. Divide by molar absorptivity of NADH to covert from absorbance (*A*) to mol/L
2. Multiply by 1,000,000 to convert from mol/L to µmol/L
3. Multiply by 3 to allow for dilution of 1 ml serum to 3 mL in the assay

Therefore:

$$1 \text{ Karmen unit} = \frac{0.001 \times 1{,}000{,}000 \times 3}{6.30 \times 10^3} \text{ U/L}$$

1 Karmen unit = 0.476 U/L

Therefore 207 Karmen units = 0.476 x 207 = **99 U/L** (2 sig figs)

b) 1 Katal/L = 1 mol/sec/L

To convert to U/L:

1. Multiply by 1,000,000 to convert from mol to µmol
2. Multiply by 60 to convert from seconds to minutes

Therefore:

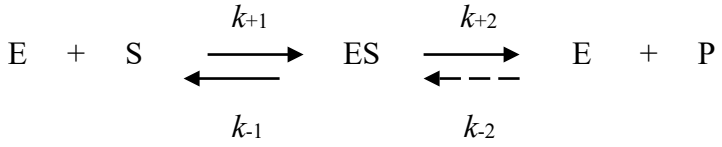
$$1 \text{ katal/L} = 1{,}000{,}000 \times 60 = 60 \times 10^6 \text{ U/L}$$

or $$1 \text{ U/L} = \frac{1}{60 \times 10^6} = 16.7 \times 10^{-9} \text{ katal/L}$$

Therefore, 99 U/L = 99 x 16.7 x 10^{-9} = 1.65 x 10^{-6} katal/L = **1.65 µkatal/L**

Effect of substrate concentration on the rate of an enzyme catalysed reaction – the Michaelis-Menten equation

Figure 9.1 listed factors which influence the rate of an enzyme catalysed reaction i.e. enzyme activity. These factors are best studied by keeping them all, except the one under investigation, constant. So far we have only considered the effect of variation of the amount of enzyme on activity since this is the variable of most interest in clinical practice. In the absence of complicating factors the rate of an enzyme reaction is directly proportional to enzyme concentration. However, the relationship between reaction rate and substrate concentration is a little more complex. If we take the simplest possible enzyme reaction in which a single substrate (S) binds to enzyme (E) to form an essential intermediate the enzyme-substrate complex (ES) which decomposes to release free enzyme and the reaction product (P) this may be represented schematically as:



The rate constants for the various reactions are denoted k_{+1}, k_{-1}, k_{+2} and k_{-2}. Note that rate constants of reverse reactions are given a minus sign. *The rate of each reaction is the product of the molar concentrations of the reactants and the respective rate constant*. Therefore we can write the following equations (in which square brackets denote molar concentrations) for the rates of formation and decomposition of the enzyme-substrate complex ES:

Rate of formation of ES $= k_{+1}[E][S] + k_{-2}[E][P]$

Rate of decomposition of ES $= k_{-1}[ES] + k_{+2}[ES]$

A few milliseconds after the enzyme and substrate are mixed [ES] builds up and does not change provided [S] is in large excess and $k_{+1} >> k_{+2}$. This condition is called a *steady state* in which the rate of formation of ES is balanced by its rate of decomposition so that [ES] *is constant.* Therefore the following *steady state equation* can be written :

Rate of formation of ES = Rate of decomposition of ES

$k_{+1}[E][S] + k_{-2}[E][P] = k_{-1}[ES] + k_{+2}[ES]$

However, in enzyme kinetics we make measurements at an early stage of the enzyme-catalyzed reaction when [P] will be very small and rate of formation of ES from E and P is very low and can be ignored. Thus if we measure initial rates only then the above equation can be simplified to:

$$k_{+1}[E][S] = k_{-1}[ES] + k_{-2}[ES]$$

then rearranged to give an expression for [E]/[ES]:

$$k_{+1}[E][S] = [ES](k_{-1} + k_{-2})$$

$$\frac{[E]}{[ES]} = \frac{(k_{-1} + k_{-2})}{k_{+1}[S]}$$

A constant, called the *Michaelis-Menten constant* (K_m) can be defined as:

$$K_m = \frac{k_{-1} + k_{-2}}{k_{+1}}$$

So substituting K_m for $(k_{-1} + k_{-2})/k_{+1}$ gives:

$$\frac{[E]}{[ES]} = \frac{K_m}{[S]} \quad \text{..... Eq. 9.1}$$

The total enzyme concentration $[E]_{Total}$ is the sum of the free enzyme and the enzyme-substrate complex. Therefore we can write the following *enzyme conservation equation*:

$$[E]_{Total} = [E] + [ES]$$

Re-arranging gives $[E] = [E]_{Total} - [ES]$, so that an expression for [E]/[ES] can be written in terms of concentrations of enzyme components:

$$\frac{[E]}{[ES]} = \frac{[E]_{Total} - [ES]}{[ES]}$$

which can also be written:

$$\frac{[E]}{[ES]} = \frac{[E]_{Total}}{[ES]} - 1 \quad \text{..... Eq. 9.2}$$

Combining the two expressions for [E]/[ES] gives:

$$\frac{[E]_{Total}}{[ES]} - 1 = \frac{K_m}{[S]}$$

Multiplication throughout by [S] gives:

$$\frac{[E]_{Total}[S]}{[ES]} - [S] = K_m$$

$$\frac{[E]_{Total}[S]}{[ES]} = K_m + [S]$$

which can be re-arranged to give an expression for [ES]:

$$[ES] = \frac{[E]_{Total}[S]}{K_m + [S]}$$

The *rate of formation of product* (v) can be expressed as a function of the rate constant k_{+2} and the concentration of the enzyme-substrate complex:

$$\boldsymbol{v = k_{+2}[ES]} \quad \text{..... Eq. 9.3}$$

Substitution of the expression for [ES] gives:

$$v = \frac{k_{+2}[E]_{Total}[S]}{K_m + [S]}$$

K_{+2} and $[E]_{Total}$ are constant and can be replaced by a single constant called *maximal velocity* (V_{max}) to give the *Michaelis-Menten equation*:

$$\boldsymbol{v = \frac{V_{max}[S]}{K_m + [S]}} \quad \text{..... Eq. 9.4}$$

Consider the behaviour of this equation at the two extremes of substrate concentration. When [S] is very low, for example much lower than the K_m (i.e. $K_m >>> [S]$), then the Michaelis-Menten equation approximates to $v = V_{max}[S]/K_m$ which is of the general form v = constant x [S]. This is a linear expression so under these conditions the rate is *directly proportional* to substrate concentration as illustrated in Fig 9.5. When a reaction

rate is proportional to a single concentration term it is said to follow *first-order kinetics*. At very high substrate concentrations most of the enzyme exists as the enzyme-substrate complex i.e. the enzyme becomes *saturated* with substrate. Under these conditions [S] >>> K_m so that the Michaelis-Menten equation approximates to $v = V_{max}[S]/[S] = V_{max}$ and the rate becomes essentially independent of substrate concentration and constant (i.e. $v = V_{max}$). Note that during derivation of the Michaelis-Menten equation V_{max} contained the term $[E]_{Total}$ so that although the rate is independent of substrate concentration it is still dependent on the amount of enzyme present. When a reaction rate is independent of concentration it is said to follow *zero-order kinetics*. This is also illustrated in Fig 9.5. At intermediate concentrations the reaction follows a mixture of first and zero-order kinetics and the plot of rate versus substrate concentration is curved. In fact it is hyperbolic with the rate approaching V_{max} asymptotically as [S] approaches infinity.

One special situation worth considering is the significance of the substrate concentration at half-maximal velocity (i.e. when $v = V_{max}/2$). We can then write:

$$\frac{V_{max}}{2} = \frac{Vmax\,[S]}{K_m + [S]}$$

which can be re-arranged to give an expression for [S]:

$$[S] = \frac{V_{max}\,(K_m + [S])}{2\,V_{max}}$$

Cancelling the V_{max} terms and re-arranging gives:

$$2\,[S] = K_m + [S]$$

and subtracting [S] from both sides gives:

$$[S] = K_m$$

Therefore the K_m is the substrate concentration at half-maximal velocity (it therefore follows that the units for K_m are concentration units) This is probably the most useful definition of K_m since it makes no assumption about the relative magnitude of the individual rate constants. If $k_{+1} >> k_{+2}$ then k_{+2} can be ignored, the ES complex is in *equilibrium* with free enzyme and substrate and K_m *approximates* to the dissociation constant of the enzyme-substrate complex:

$$K_m \approx \frac{k_{+1}}{k_{-1}} = \frac{[E][S]}{[ES]}$$

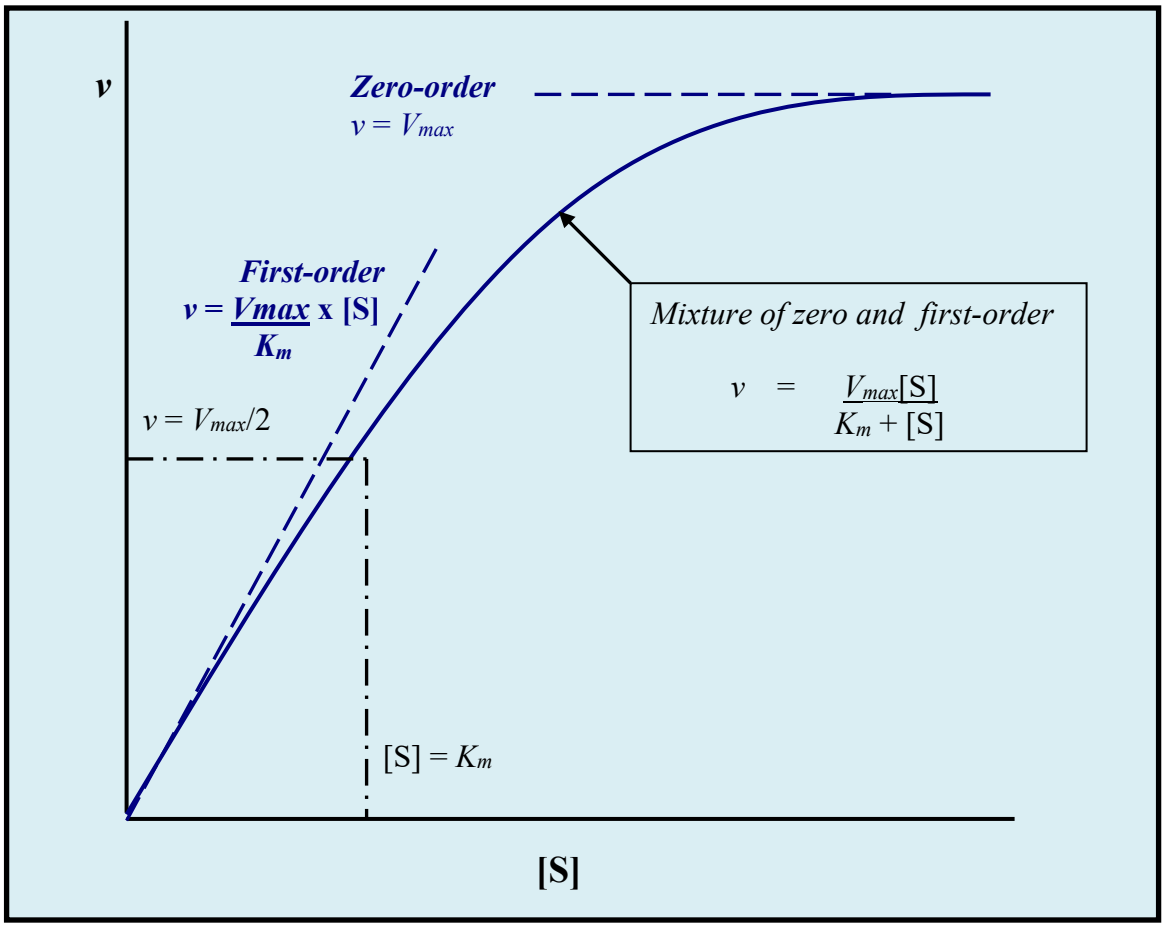


Figure 9.5 **Effect of increasing substrate concentration [S] upon the initial velocity (v) of an enzyme-catalysed reaction. At low substrate concentrations the reaction follows first order kinetics; at high substrate concentrations first order kinetcs. At all concentrations the rate is described by the Michaelis-Menten equation. At half maximal velocity the substrate concentration is equal to the Michelis-Menten constant (K_m)**

Question Q 9(3)

Enzymologists recommend that whenever possible the substrate concentration in an enzyme assay should be at least ten times the Michaelis constant (K_m). What is the rate of reaction achieved (expressed as multiples of the maximal velocity), for an enzyme reaction which obeys simple Michaelis-Menten kinetics, when the substrate concentration is exactly ten times the K_m value?

Answer Q 9(3)

The Michaelis-Menten equation relating initial velocity to substrate concentration is:

$$v = \frac{V_{max}\,[s]}{K_m + [s]}$$

v = initial velocity
V_{max} = maximal velocity (at infinite substrate concentration)
K_m = Michaelis-Menten constant = substrate concentration at half-maximal velocity
[s] = initial molar substrate concentration

Substituting $10\,K_m$ for [s]:

$$v = \frac{V_{max}\,10\,K_m}{K_m + 10\,K_m}$$

Substituting $(K_m + 10\,K_m) = 11\,K_m$, then cancelling K_m gives the value of v:

$$v = \frac{V_{max}\,10\,K_m}{11\,K_m} = \frac{10\,V_{max}}{11} = \mathbf{0.91\ V_{max}}\ (2\text{ sig figs})$$

Therefore the initial rate is approximately 90 per cent of the maximal rate (V_{max}).

Although the Michaelis-Menten equation has been derived for the simplest case of a single substrate reaction its application is by no means limited to this. The equation can be written in a more general form:

$$\boldsymbol{v = \frac{V_{max}^{app}\,[S]}{K_m^{app} + [S]}} \quad \textbf{..... Eq. 9.5}$$

in which K_m^{app} and V_{max}^{app} are not true constants but apparent values for K_m and V_{max} which depend on the concentrations of activators, inhibitors, second substrates etc which are held constant whilst [S] is varied. Variation of these apparent constants with other parameters depends on the kinetic mechanisms involved. Therefore simple modification of the Michaelis-Menten equation can be used to study enzyme inhibition, activation, pH effects and multiple-substrate reactions.

Graphical solutions of the Michaelis-Menten equation

In theory it possible to obtain estimates of K_m and V_{max} from plots of v versus [S] but deciding when the rate is maximal is difficult since the rate approaches V_{max} asymptotically. To overcome this practical difficulty various graphical solutions have been proposed.

The double-reciprocal plot of Lineweaver-Burk. Inversion of the Michaelis-Menten equation (Eq 9.4) gives the expression:

$$\frac{1}{v} = \frac{K_m + [S]}{V_{max}[S]}$$

Separating the left hand side into two components gives:

$$\frac{1}{v} = \frac{K_m}{V_{max}[S]} + \frac{[S]}{V_{max}[S]}$$

If the [S] terms are cancelled in the second component of this expression on the right hand side then this equation can be rewritten in a more useful form:

$$\frac{\mathbf{1}}{\boldsymbol{v}} = \left(\frac{\boldsymbol{K_m}}{\boldsymbol{V_{max}}} \times \frac{\mathbf{1}}{\mathbf{[S]}}\right) + \frac{\mathbf{1}}{\boldsymbol{V_{max}}} \quad \text{..... Eq. 9.6}$$

Therefore a plot of $1/v$ versus $1/[S]$ is linear, with slope K_m/V_{max} and intercept on the $1/v$ axis of $1/V_{max}$ (see Fig 9.6a). It can easily be shown that the intercept on the $1/[S]$ axis is $-1/K_m$. Double reciprocal plots are easy to interpret and computation of K_m and V_{max} is straightforward.

The [S]/v versus [S} plot of Hanes. If the double-reciprocal equation of Lineweaver and Burk (Eq 9.6) is multiplied throughout by [S] then another linear from is obtained:

$$\frac{\mathbf{[S]}}{\boldsymbol{v}} = \frac{\boldsymbol{K_m}}{\boldsymbol{V_{max}}} + \left(\mathbf{[S]} \times \frac{\mathbf{1}}{\boldsymbol{V_{max}}}\right) \quad \text{..... Eq. 9.7}$$

Thus a plot of $[S]/v$ versus [S] is linear with slope $1/V_{max}$ and intercepts on the $[S]/v$ and [S] axes of K_m/V_{max} and $-K_m$ respectively (Fig 9.6b).

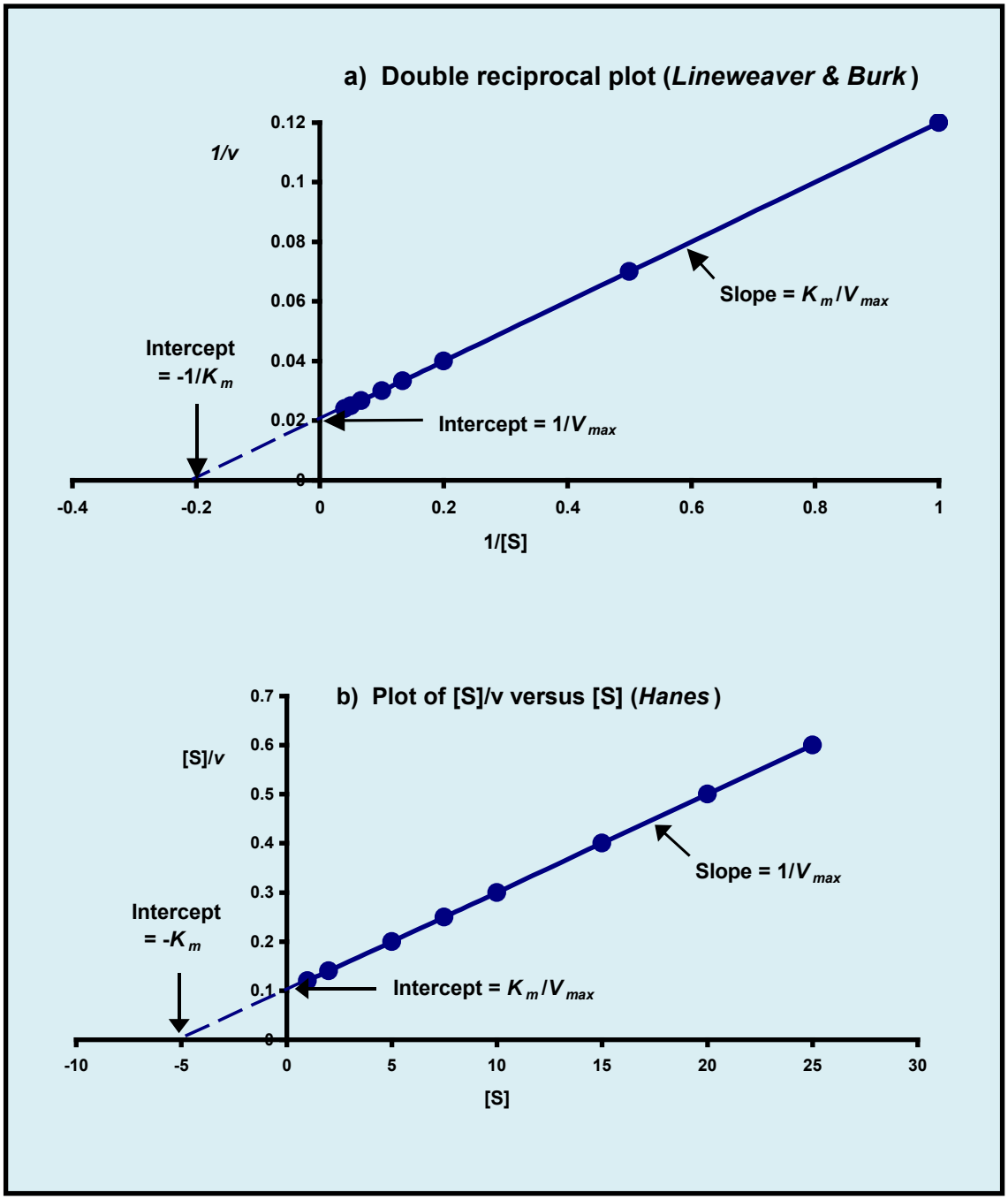


Figure 9.6 **Double reciprocal plot (a) and [S]/*v* versus [S] plot (b) for the same set of data**

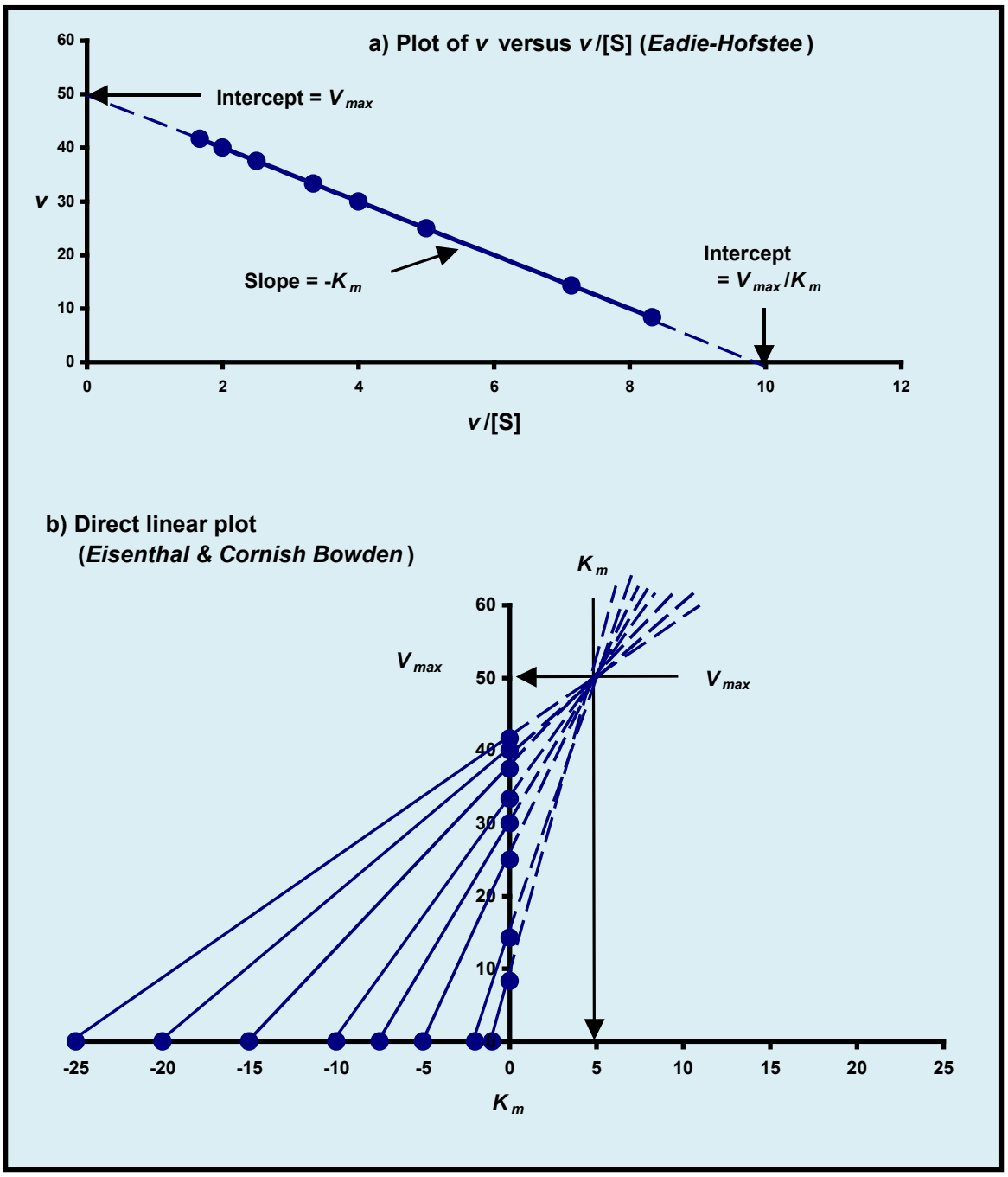


Figure 9.7 Plot of *v* versus *v*/[S] (a) and the direct linear plot (b) for the same data depicted in Fig 9.6

The v versus v/[S] plot of Eadie-Hofstee. Division of both sides of the Michaelis-Menten equation by [S] gives:

$$\frac{v}{[S]} = \frac{V_{max}}{K_m + [S]}$$

Multiplication of both sides by (Km + [S]) yields the following expression:

$$\frac{v(K_m + [S])}{[S]} = V_{max}$$

which can be simplified to:

$$\frac{vK_m}{[S]} + v = V_{Max}$$

then rearranged to give a linear expression for v and v/[S]

$$\boldsymbol{v = V_{max} - Km\left(\frac{v}{[S]}\right)} \quad \textbf{……. Eq. 9.8}$$

Therefore a plot of v versus v/[S] is linear with slope $-K_m$ and v and v/[S] intercepts of V_{max} and V_{max}/K_m respectively (Fig 9.7a).

The direct linear plot of Eisenthal and Cornish-Bowden. These authors use a rather unique approach in which the constants V_{max} and K_m are treated as variables and plotted as points in *observational space* whereas values for v and [S] are treated as constants and plotted as lines in *parameter space*. This strange concept is probably easier to understand if the Eadie-Hofstee form of the Michaelis-Menten equation (Eq. 9.5) is re-arranged slightly to:

$$\boldsymbol{V_{max} = \left(\frac{v}{[S]}\right) K_m + v} \quad \textbf{…….. Eq. 9.9}$$

This equation is now that of a straight line of the form $y = ax + b$, in which the variables x and y are now K_m and V_{max} respectively with a lope of v/[S] and y intercept of v. We are used to thinking of K_m and V_{max} being constant with an infinite number of values of v and [S] which can satisfy the Michaelis-Menten equation. This concept is now reversed with v and [S] being constant but with an infinite number of values for K_m and V_{max} which can satisfy the equation. As before we can calculate the values for the intercepts on the x and y axes when V_{max} is plotted against K_m:

When $K_m = 0$, $(v/[S])K_m$ is also zero so that Eq 9.6 becomes: $Vmax = v$. In other words the y intercept is v.

Similarly when $V_{max} = 0$, Eq 9.6 becomes: $0 = \left(\frac{v}{[S]}\right) K_m + v$

which can be re-arranged to $v = - K_m \left(\frac{v}{[S]}\right)$

Multiplying both sides by [S]/v and gives: $K_m = -[S]$. Therefore the intercept on the x axis is –[S].

This means that a line joining a pair of values for v and [S] plotted *on the y and x axes* respectively if extrapolated must pass through the point in the V_{max}- K_m space with coordinates V_{max} and K_m for that enzyme. A similar argument applies for other pairs of values for v and [S] for the same enzyme. They must all pass through the point with coordinates V_{max} and K_m. In other words they must all *intercept* at the same point and this point has the coordinates V_{max} and K_m.

The procedure is illustrated in Fig 9.7b for the same data as that used in Figs 9.6a and b and 9.7a. The values for v and [S] are marked on the V_{max} and K_m axes respectively, then each pair of values joined by a straight line which is extrapolated into the positive V_{max}-K_m quadrant. All the lines intercept at a common point with coordinates V_{max} and K_m.

Difficulties often occur in deciding upon the exact intersection point. Frequently several such points occur as a result of experimental variation (this is analogous to deciding where to draw a straight line through a series of points on a conventional plot). In the worse-case scenario if there are n pairs of observations then there is a maximum of $0.5n(n - 1)$ possible intersections. The authors recommend that the *median* of all possible intersections are used (if three lines intersect at one point the this is treated as three intersections rather than one in finding the median, if there are four lines intersecting then this counts as six intersections etc).

Although the Lineweaver-Burk plot is widely used for kinetic analysis of enzyme reactions, the use of reciprocals means that small experimental errors can result in large errors in graphically determined values for K_m and V_{max}. It has been argued that the Eadie-Hofstee plot (i.e. v versus v/[S]) results in less error. It is apparent by inspection of the plots in Figs 9.6 and 9.7, which are all based on the same v and [S] data, that each method of plotting results in a different "spread" of results e.g. the double reciprocal plot compresses the points at high substrate concentrations. This pitfall can be overcome by careful selection of concentration values employed in the experiment. The direct linear plot overcomes some of these limitations. Nowadays computer statistics packages are frequently used to fit data directly to the Michaelis-Menten equation.

Enzyme inhibition

Some enzyme inhibitors act *irreversibly* by forming a covalent bond with an amino acid residue in the enzyme thereby rendering it inactive. However, most inhibitors bind *reversibly* to the enzyme so that an equilibrium is established between the free enzyme (E) and inhibitor enzyme (EI) complex, the position of which is determined by the inhibitor constant (K_i) which reflects the affinity of the inhibitor (I) for the free enzyme:

$$E + I \rightleftharpoons EI \qquad K_i = \frac{[E][I]}{[EI]}$$

Clearly the inhibitor can lower the reaction velocity (v) by either decreasing the numerator of the Michaelis-Menten equation (i.e. value of the V_{max}^{app}) or increasing the denominator (i.e. the value for K_m^{app}). There are three main types of inhibitors which exert their effects on v in different ways:

Competitive inhibitors, often structurally related to the natural substrate bind reversibly with the enzyme at or near the active site. The inhibitor and substrate therefore compete for the enzyme according to the scheme:

$$\begin{array}{ccccccccc} E & + & S & \rightleftharpoons & ES & \rightarrow & E & + & P \\ + & & & & & & & & \\ I & & & & & & & & \\ \downarrow\uparrow & & & & & & & & \\ EI & & & & & & & & \end{array}$$

Therefore the inhibitor effectively removes a portion of the enzyme from the reaction. The enzyme conservation equation therefore becomes:

$$[E]_{Total} = [E] + [ES] + [EI]$$

By substituting $[E][I]/K_i$ for [EI] and grouping the [E] terms this becomes:

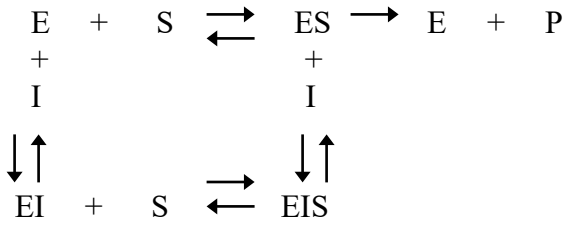
$$[E]_{Total} = [ES] + [E](1 + [I]/K_i)$$

which can the be incorporated into the steady state equation for [ES] (Eq 9.1) to derive the following variation of the Michaelis-Menten equation for competitive inhibition:

$$v = \frac{V_{max}[S]}{K_m(1 + [I]/K_i) + [S]} \quad \text{..... Eq. 9.10}$$

Therefore the K_m has been increased since it is now multiplied by a number which is always greater then one, which in turn reduces the value of v. Note that it is possible to reduce the relative contribution of the $Km(1 + [I]/K_i)$ term by increasing the value of [S]. In other words it is theoretically possible to overcome competitive inhibition at high substrate concentrations.

Non-competitive inhibitors bind reversibly at areas other than the active site. Binding of inhibitor and substrate is independent so it is possible not only to form complexes between enzyme and inhibitor (i.e. EI) but between inhibitor, substrate and enzyme (i.e. EIS) according to the scheme:



The EIS complex cannot breakdown to reaction products so again a proportion of the enzyme becomes unavailable to take part in the reaction. Inhibition cannot be overcome by increasing substrate concentration since binding at the two sites in independent. In fact the dissociation constants (K_iS) of the EIS and EI complexes are identical:

$$K_i = \frac{[E][I]}{[EI]} = \frac{[ES][I]}{[EIS]}$$

and the enzyme conservation equation becomes:

$$[E]_{Total} = [E] + [ES] + [EI] + [EIS]$$

By substituting $[E][I]/K_i$ for [EI], $[ES][I]/K_i$ for [EIS] and grouping the [E] and [ES] terms this becomes:

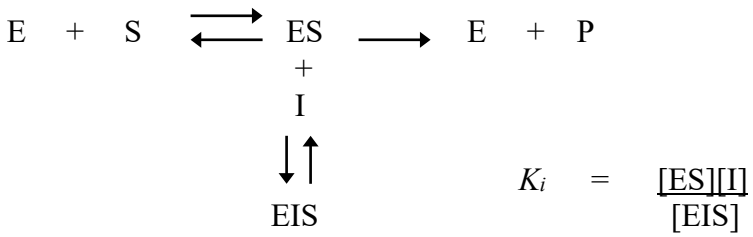
$$[E]_{Total} = [E](1 + [I]/K_i) + [ES](1 + [I]/K_i)$$

which can the be incorporated into the steady state equation for [ES](Eq 9.1) then used to derive the following variation of the Michaelis-Menten equation for non-competitive inhibition:

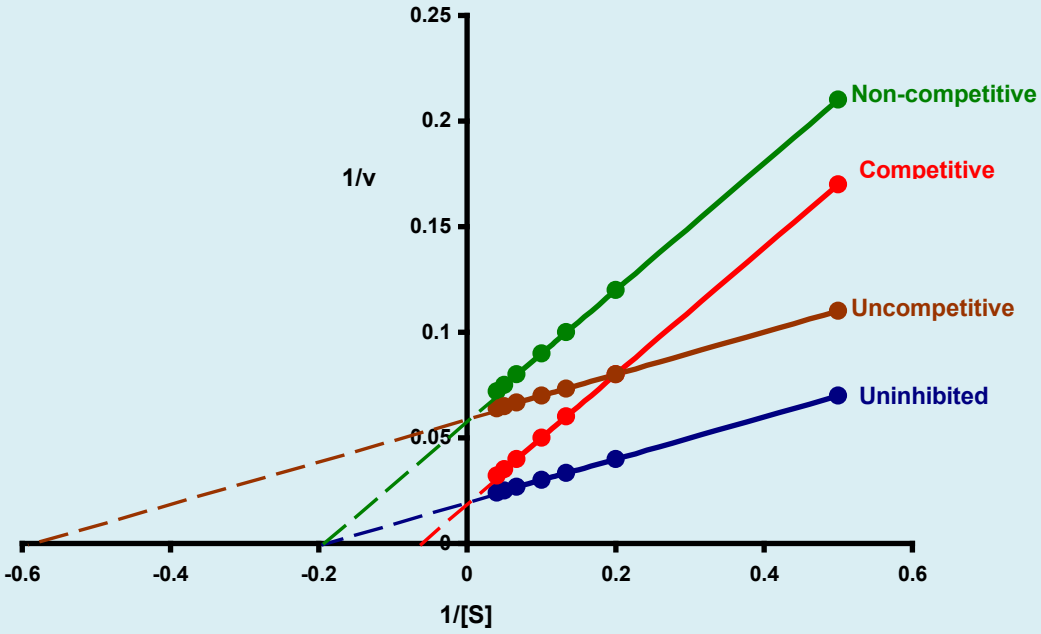
$$\boldsymbol{v = \frac{V_{max}[S]}{(1 + [I]/K_i)(K_m + [S])}} \quad \textbf{..... Eq. 9.11}$$

Therefore the value of K_m is unchanged but the V_{max} is divided by a number greater than one so its value and hence the rate of the reaction is reduced.

An *uncompetitive inhibitor* combines *only* with the ES complex, not the free enzyme according to the scheme:



$$K_i = \frac{[ES][I]}{[EIS]}$$



Type of inhibition	K_m^{app}	V_{max}^{app}
Competitive	$K_m(1 + [I]/K_i)$	V_{max}
Non-competitive	K_m	$V_{max}/(1 + [I]/K_i)$
Uncompetitive	$K_m/(1 + [I]/K_i)$	$V_{max}/(1 + [I]/K_i)$

Figure 9.8 **Double reciprocal plots for an enzyme illustrating the effects of competitive, non-competitive and uncompetitive inhibition. The inset shows the effects on the apparent K_m and V_{max}. For each plot K_m = 5 mmol/L, V_{max} = 50 µmol/min, K_i = 10 mmol/L and the inhibitor concentration is 20 mmol/L**

The enzyme conservation equation is therefore:

$$[E]_{Total} = [E] + [ES] + [EIS]$$

Substituting [EIS] = [ES][I]/K_i and grouping the [ES] terms this becomes:

$$[E]_{Total} = [E] + [ES]\,(1 + [I]/K_i)$$

which can be incorporated into the steady state equation for [ES] (Eq 9.1) then used to drive the following variation of the Michaelis-menten equation for uncompetitive inhibition:

$$v = \frac{V_{max}[S]}{K_m + [S](1 + [I]/K_i)} \qquad \text{..... Eq. 9.9}$$

As for non-competitive inhibition the V_{max} is divided by a number greater than one so its value and hence the rate of the reaction is reduced. It is interesting to note that the apparent K_m is actually reduced suggesting that the substrate is more avidly bound to the enzyme in the presence of inhibitor. However this is insufficient to overcome the reduction in apparent V_{max}.

Inhibition data can be transformed to linear equations in the same way as for uninhibited reactions. Figure 9.8 shows a double reciprocal plot for an enzyme inhibited in three different ways by the same inhibitor concentration. Simple inspection of the curves allows identification of the mode of inhibition. Values for K_m^{app} and V_{max}^{app} can be obtained by analogous procedures used for K_m and V_{max}. Substitution of values for K_m and V_{max} into K_m^{app} and V_{max}^{app} respectively, together with inhibitor concentration permits calculation of K_i. Another approach attributed to Dixon is to plot $1/v$ versus [I] using two or more substrate concentrations; the best fit lines at each substrate concentration should intercept at a value of [I] equal to K_i. The preferred approach is to determine values for K_m^{app} and V_{max}^{app} over a wide range of inhibitor concentrations then to use secondary plots to determine K_i.

Question Q 9(4)

Drug A is routinely used in the treatment of rheumatoid arthritis. It is metabolized *in vivo* to its active metabolite B by an enzyme PP. The possibility of introducing drug C into the treatment regimen is being investigated and there are some concerns that drug C may inhibit the metabolism of drug A.

8 cuvettes were set up (numbered 1 to 8) each with an optical path length of 1 cm. 0.6 mL of a stock solution of drug A (5 mmol/L) was diluted to 25 mL with buffer, then used to prepare a series of dilutions from this diluted substrate as follows:

0.5 mL of diluted A was pipetted into cuvettes 1 and 2.
0.5 mL of buffer was added to cuvette 2, mixed, then 0.5 mL transferred to cuvette 3.
0.5 mL of buffer was added to cuvette 3, mixed, then 0.5 mL transferred to cuvette 4.
0.5 mL of buffer was added to cuvette 4, mixed, then 0.5 mL removed and discarded.

0.5 mL of buffer was then added to each cuvette.

An identical set of dilutions of diluted A was prepared in cuvettes 5 to 8, except that 0.5 mL of a solution of drug C (50 mmol/L) was added at the final stage instead of buffer.

1 mL of enzyme PP solution and 1 mL of a second enzyme (which was not rate limiting but converts B into a coloured product with a molar absorptivity at 505 nm of 5500 L/mol/cm) was added to each cuvette, the contents mixed then incubated for exactly 5 minutes. The absorbance of each cuvette at 505 nm was measured versus a cuvette containing distilled water (there was no significant reagent blank).

The following absorbance readings were obtained:

Cuvette No	1	2	3	4	5	6	7	8
Absorbance	0.400	0.330	0.250	0.167	0.330	0.250	0.167	0.100

a) Calculate the **initial** concentration of substrate (A) in each cuvette
b) Calculate the **rate** of formation of B (expressed as umol /L/min)
c) Determine the **Km** of enzyme PP for A
d) Determine the **type of inhibition** present
e) Determine the **inhibitor constant** of C
f) Comment on the **likely consequences** of introducing drug C into the regimen for patients already receiving drug A.

Answer Q 9(4)

a) Concentration of stock A = 5 mmol/L = 0.005 mol/L

Concentration of diluted A (mol/L)

$$= \frac{\text{Concentration of stock A (mol/L) x Vol of stock A (mL)}}{\text{Volume of diluted A (mL)}}$$

$$= \frac{0.005 \times 0.6}{25} = 0.00012 \text{ mol/L} = 1.2 \times 10^{-4} \text{ mol/L}$$

Each cuvette contains:		
	Substrate	0.5 mL
	Buffer or inhibitor	0.5 mL
	Enzyme solution (PP)	1.0 mL
	Second enzyme	1.0 mL
	Total volume	3.0 mL

Therefore the concentration of substrate in cuvettes 1 (and 5) is given by:

Initial substrate concentration (mol/L)

$$= \frac{\text{Concentration of added substrate (mol/L) x Vol substrate added (mL)}}{\text{Final volume in cuvette (mL)}}$$

$$= \frac{1.2 \times 10^{-4} \times 0.5}{3.0} = 0.20 \times 10^{-4} = 20.0 \times 10^{-6} \text{ mol/L}$$

Since doubling dilutions were prepared the concentration in each subsequent cuvette is reduced by one half of the concentration in the previous cuvette. Therefore the concentrations are:

Cuvette No:	1	2	3	4	5	6	7	8
Substrate ($\times 10^6$ mol/L)	20	10	5	2.5	20	10	5	2.5

b) Assuming the initial rate of the conversion of of A to B is rate limiting and that the absorbance of the coloured product of the second enzyme obeys the Beer-Lambert Law then the absorbance is proportional to the concentration of substrate A consumed in the reaction catalyzed by enzyme PP. Therefore:

$$A = a \times b \times c$$

Where A = absorbance at 505 nm
a = molar absorptivity = 5,500 L/mol/cm
b = cuvette path length = 1 cm
c = concentration (mol/L)

Re-arrange to give the concentration of B in the cuvette:

$$c = \frac{A}{a \times b} = \frac{A}{5{,}500 \times 1} \text{ mol/L}$$

Divide by 5 to obtain the concentration of B produced per minute (since a 5 minute reaction time was used) then multiply by 1,000,000 to convert from mol to µmol:

$$v = \frac{A \times 1{,}000{,}000}{5{,}500 \times 1 \times 5} = A \times 36.4 \text{ µmol/L/min}$$

Cuvette No:	1	2	3	4	5	6	7	8
Absorbance	0.400	0.330	0.250	0.167	0.330	0.250	0.167	0.100
v (µmol/min)	14.6	12.0	9.1	6.1	12.0	9.1	6.1	3.6

c) To determine the K_m and answer the rest of the questions some graphical presentation of the data is required. Although not ideal, the double reciprocal plot is simplest. Cuvettes 1 to 4 are without inhibitor, cuvettes 5 to 8 have the same substrate concentrations but with inhibitor present.

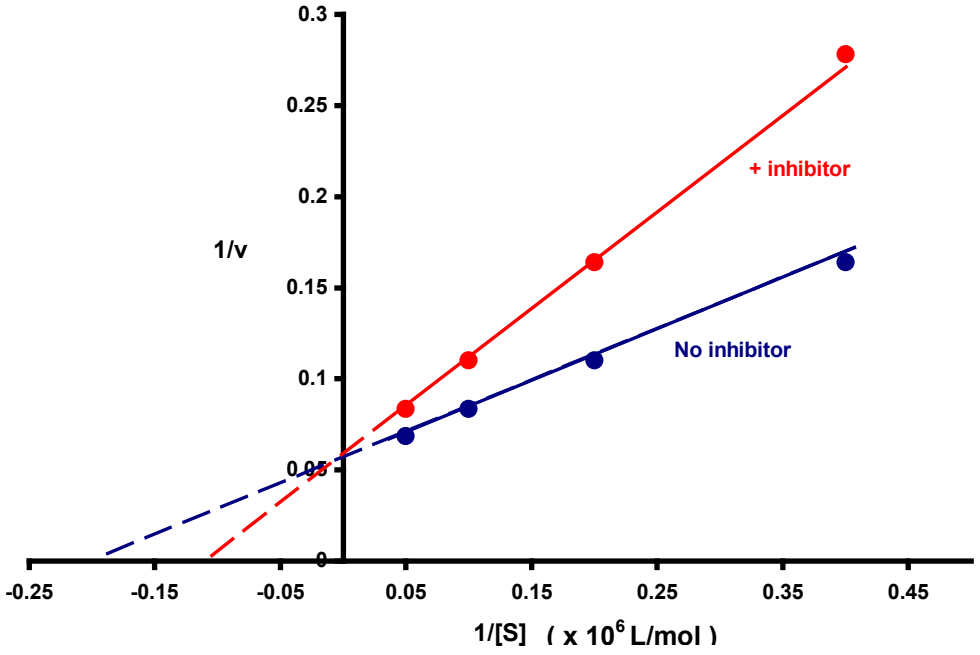
Concentration of inhibitor solution C = 50 mmol/L = 50×10^{-3} mol/L

Final concentration of inhibitor C (mol/L)

$$= \frac{\text{Initial inhibitor concentration (mol/L)} \times \text{Vol inhibitor added (mL)}}{\text{Final volume in cuvette (mL)}}$$

$$= \frac{50 \times 10^{-3} \times 0.5}{3.0} = 8.33 \times 10^{-3} \text{ mol/L}$$

Cuvette No	Inhibitor (x 10^3M)	[S] (x 10^6mol/L)	1/[S]	v	1/*v*
1	0	20	0.05	14.6	0.0685
2	0	10	0.10	12.0	0.0833
3	0	5	0.20	9.1	0.110
4	0	2.5	0.40	6.1	0.164
5	8.33	20	0.05	12.0	0.0833
6	8.33	10	0.10	9.1	0.110
7	8.33	5	0.20	6.1	0.164
8	8.33	2.5	0.40	3.6	0.278



From the graph the intercept on the 1/[S] axis for the uninhibited reaction is approximately - (0.20 x 10^6) L/mol. This corresponds to -1/K_m:

$$\frac{-1}{K_m} = -(0.20 \times 10^6)\ \text{L/mol}$$

Therefore $$K_m = \frac{-1}{-(0.20 \times 10^6)} = \mathbf{5.0 \times 10^{-6}\ mol/L}$$

d) Drug C lowers the activity of the enzyme by increasing the value of the K_m without altering the V_{max} i.e. the double reciprocal plots for the inhibited and uninhibited reaction intersect on the $1/v$ axis. Therefore drug C is a **competitive inhibitor** of enzyme PP.

e) From the graph the intercept on the 1/[S] axis for the inhibited reaction is approximately - (0.11 x 10^6) L/mol. This corresponds to $-1/K_m^{app}$:

$$\frac{-1}{K_m^{app}} = -(0.11 \times 10^6)\ \text{L/mol}$$

Therefore $$K_m^{app} = \frac{-1}{-(0.11 \times 10^6)} = 9.09 \times 10^{-6}\ \text{mol/L}$$

Since we are dealing with a competitive inhibitor:

$$K_m^{app} = K_m\,(1 + [I]/K_i) \qquad \text{(see Fig 9.8)}$$

Where K_m = K_m of uninhibited reaction = 5.0×10^{-6} mol/L
[I] = inhibitor concentration = 8.33×10^{-3} mol/L
K_i = inhibitor constant

Substituting these values and solving for K_i:

$$9.09 \times 10^{-6} = 5.0 \times 10^{-6}\,\{1 + (8.33 \times 10^{-3})/K_i\}$$

$$1 + \frac{(8.33 \times 10^{-3})}{K_i} = \frac{9.09 \times 10^{-6}}{5.0 \times 10^{-6}}$$

$$\frac{(8.33 \times 10^{-3})}{K_i} = 1.818 - 1 = 0.818$$

$$K_i = \frac{8.33 \times 10^{-3}}{0.818} = \mathbf{1.0 \times 10^{-2}\ mol/L}\ (2\ \text{sig figs})$$

f) The inhibitor constant (K_i) of drug C is considerably higher than the K_m for the substrate A (1.0×10^{-2} mol/L compared to 5.0×10^{-6} mol/L). Both of these constants are inversely proportional to the affinity of the enzyme PP for the substrate and inhibitor. Thus the affinity of the enzyme for the substrate, A, is considerably greater than its affinity for the inhibitor, C. Therefore the effect of

introducing drug C into the regimen for patients receiving drug A will depend on the relative concentrations of the two drugs. If their therapeutic concentrations are similar, or the concentration of A is greater than C, then C will have little effect on the metabolism of drug A or the optimum dose required to achieve therapeutic levels of its active metabolite B. If, however, the plasma level of drug C is considerably higher than that of drug A, then inhibition of the metabolism of drug A will occur and higher plasma levels of drug A will be achieved, with the consequence of decreased levels of the active metabolite B. As a result higher doses of dose A will be required to achieve the same therapeutic result.

If drug C is not only an inhibitor of PP but a substrate for this enzyme then the metabolism of drug C will also be affected by drug A which will have consequences on the levels of drug C (and its metabolites) achieved.

ADDITIONAL QUESTIONS

1. An assay mixture for the measurement of lactate dehydrogenase constituted 2.7 mL of buffered NADH and 100 μL of serum. The reaction was started by adding 100 μL of sodium pyruvate. The absorbance change over 5 minutes was 0.150 when measured in a 0.5 cm light path at 340 nm. Assuming the molar absorbtivity of NADH at 340 nm is 6.30 x 10^3 $L.mol^{-1}cm^{-1}$, calculate the enzyme activity in international units per litre of serum..

2. An assay for alkaline phosphatase activity involved mixing 0.5 mL of serum with 2.7 mL buffer, allowing temperature to reach equilibrium then starting the reaction by adding 0.2 mL of substrate (4-nitrophenyl phosphate). The increase in absorbance in a 1cm cuvette due to the liberation of product (4-nitrophenol) was 0.180 over a 5-minute period. Calculate the alkaline phosphtase activity expressing the result as a) international units per litre of serum, and b) katals per litre of serum. Assume that the molar absorptivity of 4-nitrophenol is 1.88 x 10^4 L/mol/cm.

3. The Somogyi saccharogenic method for the assay of amylase involves measuring the rate of release of glucose from substrate. One Somgyi unit is the amount of enzyme catalysing the release of 1 mg of glucose in 30 min per 100 mL serum. Derive a factor to convert Somogyi units to international units per litre of serum.

4. One Wroblewski-laDue unit is the amount of lactate dehydrogenase which results in an absorbance change (due to NADH) at 340 nm of 0.001 per minute per mL serum in a reaction mixture with a total volume of 3 mL. Derive a factor to convert Wroblewski-LaDue units to International units per litre of serum. Assume the molar absorptivity of NADH is 6.3×10^3 L/mol/cm.

5. If the K_m of an enzyme which obeys simple Michaelis-Menten kinetics is 2.5 mmol/L, what velocity (expressed as a multiple of V_{max}) would be obtained at a substrate concentration of 10 mmol/L?

6. What information can be obtained from the double-reciprocal plot for an enzyme under the following conditions: a) $1/v = 0$ when $1/[S] = -12.5 \times 10^6$ L/mol, b) $1/[S] = 0$ when $1/v = 5.2 \times 10^6$ min.L/mol, c) $1/[S] = 0$ when $1/v = 6.5 \times 10^6$ min/mol and the slope of the line is 100 min/L?

7. You carry out an enzyme experiment in which the substrate concentration is expressed as mmol/L and the reaction velocity in µmol/L/min. What would be the units for the axes of the three following plots: a) 1/[S] versus $1/v$, b) $[S]/v$ versus [S], c) v versus $v/[S]$?

8. Mucic acid is an inhibitor of β-glucuronidase. The following data were obtained using phenolphthalein glucuronide as substrate, in the presence and absence of mucic acid (concentration in the assay = 1.0×10^{-4} mol/L).

Substrate Concentration (mmol/L)	Reaction velocity	
	No inhibitor	Mucic acid
0.5	33	9
1.0	50	17
2.0	67	29
4.0	80	44
10	91	67

Determine the type of inhibition and the enzyme-inhibitor dissociation constant.

9. An experiment was conducted to study the effect of pH on the activity of lactate dehydrogenase. Using a histidine buffer at pH 5.5 and 7.4 the reaction was monitored by following the increase in absorbance at 340 nm due to the reduction of NAD. The following data were obtained:

Lactate concentration mmol/L	Reaction velocity	
	pH 7.4	pH 5.5
1	12	33
2	21	50
4	35	67
10	57	83
20	73	91

Stating any assumptions that you make determine the pH at which the enzyme has greatest affinity for the substrate.

10. The apparent K_m and V_{max} of an enzyme were measured over a range of inhibitor concentrations and the following data obtained:

Inhibitor concentration (mmol/L)	Apparent value	
	K_m (mmol/L)	V_{max} (μmol/min)
5	10	7.5
10	7	5
15	5	4
20	4	3

Determine the mode of inhibition and the inhibitor constant (K_i).

Chapter 10

The basis of statistics

"If your experiment needs staticism then you ought to have done a better experiment"

In the perfect world if we were to measure the creatinine concentration in a serum sample a large number of times then we would obtain exactly the same result every time. In practice, however, the result is not always the same due to *analytical imprecision*. The results obtained would vary but only by a small amount, often the same result would be obtained more than once so that the results would tend to cluster around a particular value. The value around which results would cluster is not necessarily the true value due to inherent *inaccuracy* of the method. Similarly if serum samples were collected from the same "normal" individual on a number of occasions and the creatinine concentration measured in each sample then a similar cluster of results would be obtained but the spread would be much wider due to *intra-individual variation* being added to the analytical imprecision. On the other hand if samples were collected from a number of "normal" individuals then the spread of results would be even wider due to a contribution from *inter-individual variation* in addition to intra-individual variation and analytical imprecision. An important consequence is that if two different results are obtained we cannot be sure that the change is real since it may be explained by the expected analytical impression and/or intra-individual variation.

The science of statistics gives us the tools to deal with this *random* variation due to analytical imprecision and biological variation in order that we can extract maximum information from data that we obtain in the clinical laboratory. Nowadays anyone can use computers (and some pocket calculators) to do statistical calculations. However, correct interpretation of the statistical parameters produced requires some understanding of the underlying principles.

The presentation and description of laboratory data

Figure 10.1a gives the results obtained when creatinine concentration was measured in sera from 60 "normal" individuals. Simple inspection of the data shows that:

- The results are not all the same
- All the results fall within the range 35 to 122 μmol/L
- The same result often appears more than once

In most areas of statistics an extremely useful first step is to express the results in the form of a diagram. Depending upon the number of results the data is first grouped into class intervals of equal size. The interval used should be chosen to make sure that most intervals contain more than one result. For the creatinine data in Fig 10.1 a class interval spanning a concentration range of 10 μmol/L ensures that each group contains at least one result with one of the groups containing as many as sixteen values (Fig 10.1b). If a graph is plotted with the concentration intervals as the *x* axis and frequency as the *y* axis then the result (Fig 10.1cc) is a *frequency distribution* or *histogram*.

Visual inspection of the distribution reveals that the class interval with the highest number of results is the 70-79 μmol/L group and that there are approximately equal numbers of results below this group as above it i.e. the overall shape is *symmetrical*. As we move further away from this group on either side of the diagram the number of results in each group diminishes. If we were to join up the peaks of each class interval then the result would be a *continuous bell-shaped curve* which mathematicians refer to as a *normal* or *Gaussian* distribution.

It is often useful to find a mathematical way to describe this distribution or curve. We need to convey two things:

1. The peak value. Mathematicians call this the *measure of central location*.
2. The spread of results (i.e. a measure of the variability of the results) or a measure of the width of the bell-shaped curve. Mathematicians call this the *measure of dispersion*.

a) Individual creatinine results (µmol/L)

80	62	91	81	111	75	78	45	91	72	65	50	57	85	103	75	75
63	47	93	77	122	60	76	97	35	54	106	90	80	80	74	41	78
76	115	76	52	87	75	82	64	93	71	79	100	67	68	74	59	69
63	68	65	89	72	82	68	80	83								

b) Frequency of results grouped in intervals spanning 10 µmol/L

Interval:	**30-39**	**40-49**	**50-59**	**60-69**	**70-79**	**80-89**	**90-99**	**100-109**	**110-119**	**102-129**
Frequency:	**1**	**3**	**5**	**12**	**16**	**11**	**6**	**3**	**2**	**1**

c) Histogram of frequency versus concentration interval

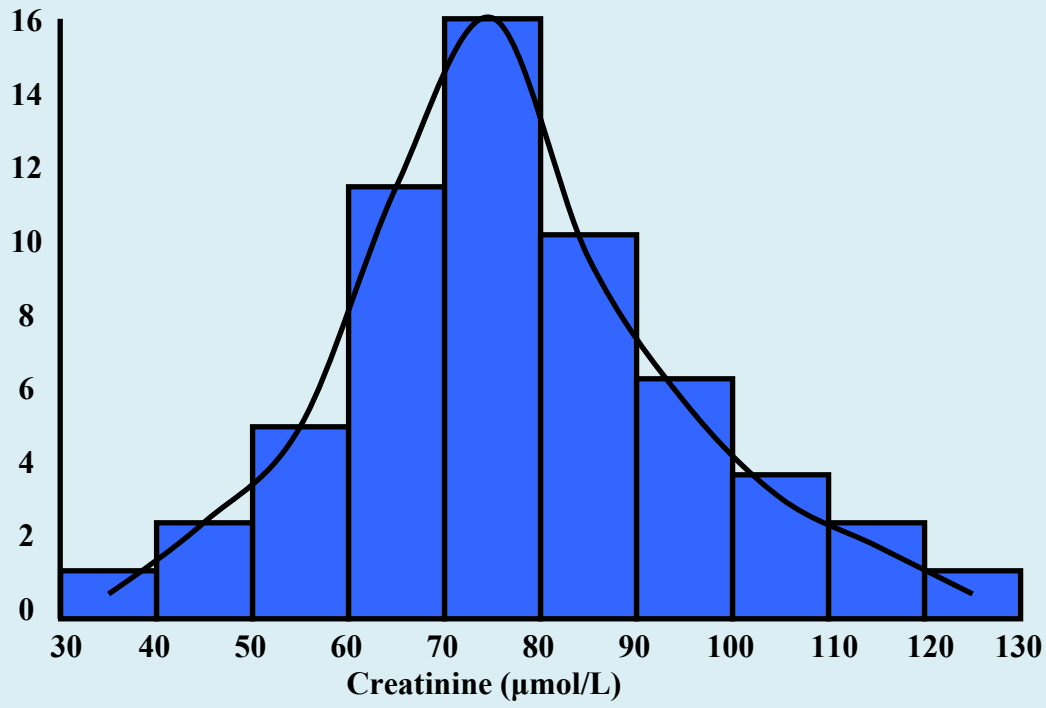


Figure 10.1 A set of 60 serum creatinine results obtained on healthy individuals (a), grouped into concentration intervals (b) and plotted as a histogram (c)

Measures of central location

The arithmetic mean or average: This simply is the sum of all the individual results in the series (if the same result is encountered more than once then it is counted more than once), divided by the number of results. If we denote each individual result in the series by the symbol x (so that the first is x_1, the second x_2, etc), the symbol Σ to mean the sum of all values of the series and n as the number of results, then we can write the following expression for the mean (m):

$$\textbf{Mean } (m) \quad = \quad \frac{\Sigma x}{n} \quad \text{..... Eq. 10.1}$$

The median: This is simply the value such that half of the data points fall above it and a half below it. In other words if we have 100 results and arrange then in ascending order, then the value of the fiftieth result is the median.

The mode: The mode is the most frequently occurring result or the class interval containing the most results.

Measures of dispersion

The standard deviation (*SD* or *s*): This is a measure of the average difference of all the values from the mean. If the mean result (m) is subtracted from an individual result (x) then the result is the *difference* or *deviation* of that result from the mean i.e. ($x - m$). If this is done for each data point (i.e. each individual result) and these deviations are added together, then the result can be expressed mathematically as $\Sigma (x - m)$. If this value is divided by the number of results, n, then the result would be expected to be a measure of the average difference of all the results from the mean. However, this is not the case, the result comes out at zero. The reason for this is that the normal distribution is symmetrical with an approximately equal number of results both below and above the mean. The deviation of a result below the mean is negative, the deviation of a result above the mean is positive. Therefore the positive deviations cancel the negative deviations so that their sum is zero. To overcome this problem mathematicians *square* each deviation so as to always give a positive result. A positive deviation multiplied by a positive deviation gives a positive result, as does a negative deviation multiplied by a negative deviation. If these are then added together then the result is the *sum of squares* of the deviations, denoted by the expression:

$\Sigma(x - m)^2$. Mathematicians use this trick in many areas of statistics. If the sum of squares is divided by the number of data points, n, then the result is a measure of dispersion known as the *variance*, which is denoted by the symbol s^2. If the square root of the variance is taken (to allow for taking squares of the deviations in the first place), then the result is the *standard deviation*, denoted by the symbol s or SD, a value which is more easily related to the dispersion of results in the distribution. This simple concept is complicated by the fact that instead of dividing by n it is customary to divide by n-1. n-1 is known as the *degrees of freedom*. The reason for this that when the sum of deviations (or their squares) is calculated, use of the value for the mean restricts the freedom of the individual values. Suppose we had six results, 1, 2, 3, 4, 5 and 6 (the numbers on a dice). Their sum is 21 and their mean (21/6) is 3.5. The deviations $(x - m)$ for the first five values are –2.5, -1.5, -0.5, +0.5, +1.5 and their sum, $\Sigma (x - m)$, is –2.5. Since the deviations must add up to zero, it follows that the deviation for the sixth value must be +2.5. Therefore the sixth value must be the sum of the mean and its deviation i.e. 3.5 + 2.5 = 6. In other words the final value in the series is fixed, cannot vary and so does not add any useful information Therefore for practical purposes there are only 5 results contributing to the sum of squares. Another way of looking at this is if a dice is lying with the six side face down then we do not need to turn the dice over to know what the value on the hidden face is! In practice the larger the number of data points the less important the difference between the number of values (n) and their degrees of freedom (n-1); above n=30 this difference is usually ignored. The expressions for variance and standard deviation are:

$$\textbf{Variance } (s^2) = \frac{\Sigma (x - m)^2}{n - 1} \quad \text{..... Eq. 10.2}$$

$$\textbf{Standard deviation } (s \text{ or } SD) = \sqrt{s^2} = \sqrt{\frac{\Sigma (x - m)^2}{n - 1}} \quad \text{..... Eq. 10.3}$$

The units of standard deviation are the same as the units of the data used in its calculation. In an attempt to standardise the expression of *s,* clinical biochemists often use the term *coefficient of variation*, denoted *cv*. This is the standard deviation divided by the mean; the result is usually expressed as a percentage:

$$\text{Coefficient of variation } (\%cv) = \frac{s \times 100}{m} \quad \ldots\ldots\ldots\ldots\text{Eq 10.4}$$

The hope was that by expressing imprecision in this way, the same numerical result would be obtained over the entire concentration range of the assay. However, this is rarely the case.

The range: This is simply the difference between the highest and the lowest value in the set of data. This is the least reliable measure of dispersion.

The interquartile range: The data are arranged in ascending order and grouped into four equal sets (known as quartiles). The middle two sets (comprising the middle fifty per cent of the data points) form the interquartile range.

Is the data "normally distributed"?

Often the histogram of a set of data is not a typical bell-shaped Guassian distribution. Fig 10.2 shows two ways in which the curve may deviate from normality. In Fig 10.2a, the two curves are not symmetrical but *skewed*. Statistical packages often calculate a parameter called the *skew*. A skew of *0* indicates no skew, a positive value indicates skew to the right (curve A) and a negative value skew to the left (curve B). A skewed distribution can often be converted to a reasonably "normal" distribution by taking logarithms of the data e.g. the distribution of serum bilirubin concentrations in normal adults is normally skewed to the right whereas the distribution of the logarithm of concentration becomes relatively normal.

In Fig 10.2b the curves differ in how "peaked" or "flat" they are; this is known as *kurtosis*. Again statistical packages often calculate a value for kurtosis. A truly Guassian curve has a kurtosis of 3 (some computer programs convert this value to zero).

Formal tests of normality include:

- The Anderson Darling test
- The Kolmogorov Smirnov test
- The Shapiro Wilks W test
- The Shapiro Francia W test

A discussion of these tests is beyond the scope of this book but may be found in standard statistical texts.

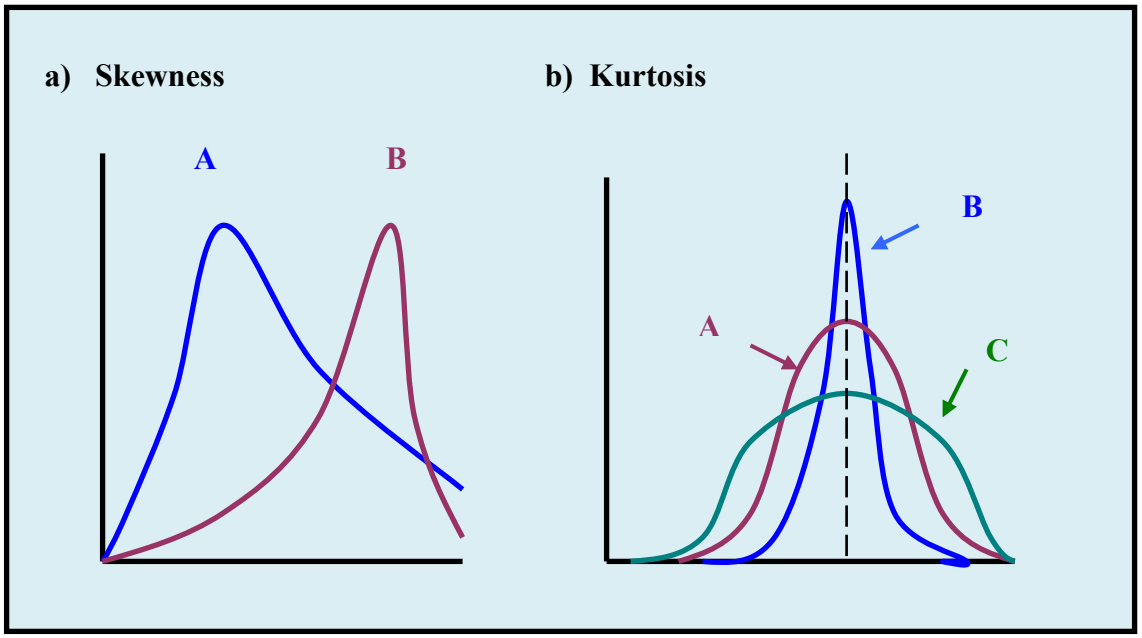


Figure 10.2 **Deviations from true Guassian distributions a) Curve A is skewed to the right (*positive skewness*), curve B is skewed to the left (*negative skewness*); b) Curve B is "normal" or *mesokurtic*, curve A is more "peaked" or *leptokurtic*, curve C is flatter than normal or *platykurtic***

Question Q 10(1)

A laboratory had just changed its method for serum creatinine. To check that there had been no change in their reference range they analysed serum samples collected from 12 members of staff and obtained the following results (arranged in ascending order): 44, 58, 60, 68, 70, 75, 76, 78, 80, 90, 95, 106 µmol/L. Calculate the mean, variance, standard deviation and coefficient of variation.

Answer Q 10(1)

Construct a table with the individual creatinine concentrations (x) in the first column:

	x	$(x - m)$	$(x - m)^2$	x^2
	44	-31	961	1936
	58	-17	289	3364
	60	-15	225	3600
	68	-7	49	4624
	70	-5	25	4900
	75	0	0	5625
	76	1	1	5776
	78	3	9	6084
	80	5	25	6400
	90	15	225	8100
	95	20	400	9025
	106	31	961	11236
Total:	**900**	**0**	**3170**	**70670**

Add these together to give their sum (Σx). The number of results (n) is 12.
From Σx and n, the mean (m) can be calculated:

$$\text{Mean } (m) = \frac{\Sigma x}{n} = \frac{900}{12} = \mathbf{75\ \mu mol/L}$$

Next subtract the mean (m) from each individual value of x so as to give a column for individual deviations ($x - m$). Note that the sum of all these, $\Sigma(x - m)$, is zero and cannot be used in the calculation of variance. Instead calculate the square of these deviations i.e. $(x - m)^2$ and enter in the third column. These are then added together to give the sum of squares of the individual deviations, $\Sigma(x - m)^2$, referred to by mathematicians as simply the *sum of squares*. This value can then be used to calculate the variance:

$$\text{Variance } (s^2) = \frac{\Sigma (x - m)^2}{n - 1} = \frac{3170}{(12 - 1)} = \frac{3170}{11} = \mathbf{288\ \mu mol/L}$$

The standard deviation is simply the square root of the variance:

$$\text{Standard deviation } (s) = \sqrt{s^2} = \sqrt{288} = \mathbf{17\ \mu mol/L}$$

The coefficient of variation is the standard deviation expressed as a percentage of the mean:

$$\text{Coefficient of variation } (cv) = \frac{s \times 100}{m} = \frac{17 \times 100}{75} = \mathbf{23\%} \text{ (2 sig figs)}$$

Before the advent of computers or sophisticated pocket calculators it was often easier to calculate the sum of the squares of individual values (Σx^2), then calculate the sum of squares of the deviations using the identity:

$$\Sigma(x-m)^2 = \Sigma x^2 - \frac{(\Sigma x)^2}{n} \quad \text{or} \quad \Sigma x^2 - n\,m^2$$

$$\text{i.e.} \quad \Sigma(x-m)^2 = 70670 - \frac{900^2}{12} = 70670 - 67500 = 3170$$

$$\text{or} \quad \Sigma(x-m)^2 = 70670 - (12 \times 75^2) = 70670 - 67500 = 3170$$

Does a single result belong to the population?

This a question we often try to answer in clinical biochemistry. The creatinine data quoted in Fig 10.1 and Question Q 10(1) were obtained from normal individuals. We may want to know if the creatinine result obtained from a patient is "abnormal" i.e. is it "different" from the reference population? There is no perfect way to answer this question. Statisticians try and deal with this problem by reformulating the question as "suppose this result does belong to this normal population what is the likelihood that this result could have been obtained by pure chance?" If it is improbable that it arose by chance then it is probably "significantly " different (in this case abnormal). This begs the question as to how unlikely the event has to be for the result to be considered different or belonging to a different population? By convention a probability of less than 1 in 20 (i.e. 0.05 or 5%) is taken as a cut-off indicating that the likelihood of the result not being abnormal as so low as to be negligible. This value was suggested by the eminent mathematician Fischer and is not based on any theoretical consideration. Statistics can only answer the question "how likely it is that an event has occurred by chance", whether the difference matters is a subjective one!

But how is this probability obtained from the individual result and the data in the reference population? The reference population when plotted (see Fig 10.1) shows a bell-shaped curve and its characteristics are partially determined by the peak value (the mean, *m*) and the width (standard deviation, *s*). Mathematicians have shown that the curve can be described mathematically by the exponential equation:

$$y = \frac{1}{s\sqrt{2\pi}} \times e^{-\frac{(x-m)^2}{2s^2}} \quad \text{..... Eq. 10.5}$$

In other words the value of *y* is a complicated function of both *m* and *s* (some forms of this equation use μ and σ instead of *m* and *s* but the subtle difference need not concern us here). Rest assured that you will *never* need to manipulate this equation. However, what we need to know is not the value of *y* but the probability of obtaining any particular value of *x*. This probability is given by the area under the curve if a perpendicular line is drawn at point *x*. Since all results for all the population must fall somewhere within the curve (probability = 1) the total area must equal 1. This area can be calculated by a complex mathematical function obtained by integrating equation Eq. 10.5. Thus from the values of *x*, *m* and *s* it is possible to calculate the probability of obtaining a value *x*; if it is less than 0.05 then the result is significantly different. It is obviously inconvenient to have to perform such a complex calculation every time. To get around this difficulty mathematicians always reduce their data to a "normalized population" in which the mean is always *zero* and the standard deviation *one*. As a result the calculation need only be done once and is used to generate a statistical table in which the probability of obtaining any value of *x* can be easily obtained. The mean is subtracted from the value then divided by the standard deviation to give a *standard score*, z:

$$z = \frac{x - m}{s} \quad \text{..... Eq. 10.6}$$

z (sometimes called the *normal deviate*, *d*, or the *standard deviation index*, *SDI*) is therefore the number of standard deviations the value of *x* is away from the mean. *z* is always normally distributed with a mean of zero and standard deviation of one. Such a normalized Guassian distribution (in which *z* is the horizontal axis) is shown in Fig 10.3. The curve has a mean of zero and a standard deviation of one. Since all results of the population must fall somewhere in the area between the curve and the *z* axis, the total area under the curve is the probability (*P*) of obtaining any value and is one.

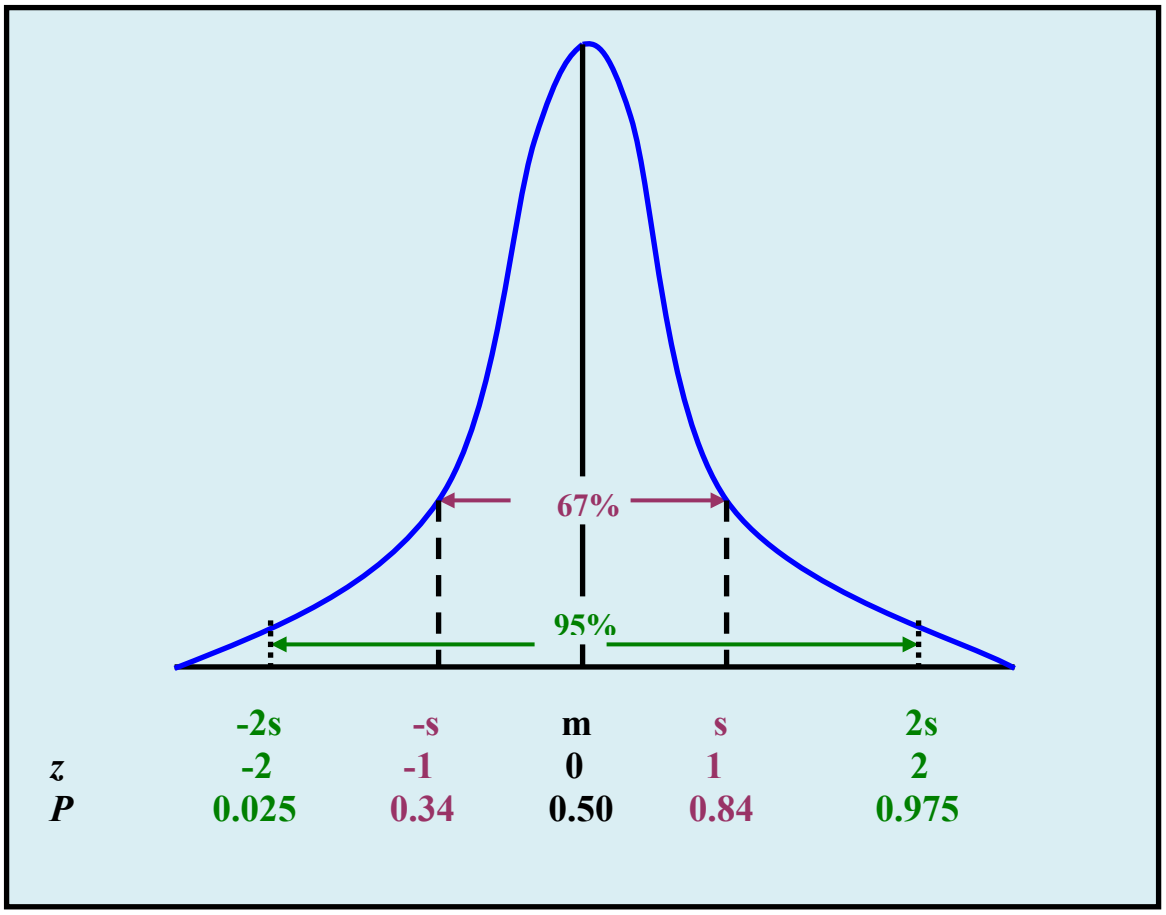


Figure 10.3 Normalised Guassian distribution in which the mean is zero and standard deviation 1. Values of P show the probability of obtaining a value to the left of the z value i.e. area under the curve to the left of a perpendicular line drawn at the value for z

The probability of obtaining a result between any two given values is the area under the curve between the two values. For example, the probability of obtaining a value between the mean minus one *s* and the mean plus one *s* is two-thirds or 0.67 (equal to 67%). The probability of obtaining a value between the mean minus 2*s* and the mean plus 2*s* is 19/20 or 0.95 (equal to 95%). Strictly speaking 1.96*s* should be used rather than 2*s*. It follows that the probability of obtaining a result outside the mean ± 2*s* (or more correctly mean ± 1.96*s*) is 1 – 0.95 = 0.05 (or 5%). This is exactly the level of probability which statisticians regard as significant when deciding whether a particular value belongs to a given population.

Since 5% of values fall within the mean ±1.96*s* and the normal curve is symmetrical it follows that 2.5% or results will be below the mean – 1.96*s* and a further 2.5% will be above the mean + 1.96*s*. This is quite an important point since in some situations we only wish to know if a result is significantly greater than the mean + 1.96s or is significantly less than the mean – 1.96s. In this case we use the value of *z* which excludes 10% of results (*P*=0.1) since only a half of these (5%) will be greater (or less than) the range encompassing 90% of the values. The value of *z* which gives rise to range which excludes the lowest and highest 5% of results is 1.645. A table of *z* values with their corresponding probabilities is given in Fig 10.4.

P	**0.33**	**0.10**	**0.05**	**0.02**	**0.01**	**0.002**	**0.001**
z	**1.0**	**1.645**	**1.96**	**2.326**	**2.576**	**3.090**	**3.291**

Figure 10.4 Percentage points of the normal distribution. This table gives the percentage points most frequently required for significance tests for a normal variable having zero mean and unit standard deviation. Thus, the probability of obtaining a departure from the mean of more than 1.96 standard deviations in either direction is 0.05 or 5%

The range of values obtained at any probability level is known as the *confidence limits*. The 95% confidence limits of a set of results can be calculated from the corresponding *z*, the mean (*m*) and standard deviation (s) as follows:

Upper limit = $m + (z \times s)$

Lower limit = $m - (z \times s)$

In general the 95% confidence limits (where $z = 1.96$) are:

Mean - (1.96 x *s*) *to* mean + (1.96 x *s*) …….. Eq. 10.7

It is common practice in clinical biochemistry to use the 95% confidence limits obtained for the concentration of an analyte from normal subjects as the *reference range*. Any value obtained for a patient outside this range is usually regarded as abnormal. The truth

is that the probability of obtaining a result outside of this range is 0.05 or 1 in 20 – we are simply making a subjective judgment that it is abnormal. If we were to measure this analyte in twenty healthy individuals then we would expect one of them to have a result outside of the reference range. Similarly, if we were to measure twenty different analytes in the same patient then again one of the results would most likely fall outside of the reference range.

Calculation from the creatinine data for 60 normal individuals depicted in Fig 10.1 gives a mean of 76 µmol/L and standard deviation of 17 µmol/L. From these figures it is possible to calculate the 95 % confidence limits (when $z = 1.96$) as follows:

95% confidence limits = mean ± (1.96 x s)
= 76 ± (1.96 x 17)
= 76 ± 33 (to 2 sig figs)
= (76 – 33) to (76 + 33)
= **43 to 109 µmol/L**

The chance of obtaining a result outside of these limits is 100 - 95 = 5% (1 in 20) and if it occurs is *probably* abnormal. The chance of obtaining a result less than 43 µmol/L is 2.5% (1 in 40), and of obtaining a result greater than 109 µmol/L is also 2.5% (1 in 40). If we wish to know the chance of obtaining a result greater than 115 µmol/L then the first thing to do it to calculate the z score:

$$z = \frac{x - m}{s} = \frac{115 - 60}{17} = \frac{55}{17} = 3.24$$

Inspection of the table in Fig 10.4 shows that there is no value of z corresponding exactly to $z = 3.24$, but that when $z = 3.09$ the probability (P value) is 0.002. Therefore the chance of obtaining a value outside the mean ± 3.24s is slightly less than 0.002. The chance of obtaining a result greater then mean + 3.24s is one half of this i.e. 0.001 (or 1 in 1000).

Question Q 10(2)

The imprecision of a certain assay for Troponin I yields a coefficient of variation of 13% between 0.3 and 0.5 µg/L, around the decision point for myocardial infarction of 0.4 µg/L. A result of 0.46 µg/L is obtained on a sample. Assuming that is the true level of Troponin I, give an estimate of the probability that analysis of that same sample would give a result below the decision point.

Answer Q10(2)

The true result (0.46 µg/L) can be considered as the mean with a coefficient of variation (cv) of 13%. The first step it to calculate the standard deviation (s). cv, mean (*m*) and s are related as shown in Eq 10.4:

$$cv\ (\%) = \frac{s \times 100}{m}$$

Which can be rearranged to:

$$s = \frac{cv\ (\%) \times m}{100}$$

Substitute for *cv* and *m* and solve for *s*:

$$s = \frac{13 \times 0.46}{100} = 0.060\ \mu g/L$$

Therefore the analyses of the sample are distributed with a mean of 0.46 µg/L and *s* of 0.06 µg/L. We want find out what proportion of results will be below the decision point of 0.4 µg/L. To do this we need to 'normalize' the data so that the mean is zero and the SD =1. i.e. calculate the standard deviate -'z':

$$z = \frac{\text{decision point} - m}{s} = \frac{0.4 - 0.46}{0.06} = \frac{-0.06}{0.06} = -1$$

Therefore the decision point is –1*s* from the mean. From the table in Fig 10.4 we can see that the probability of obtaining a result which differs from the mean by more than 1*s* in either direction (i.e. when $z = 1$ or –1) is 0.33.

Therefore the probability of obtaining a result below the decision point is one half of 0.33 i.e. 0.17 (2 sig figs).

Utilization of the Normal Distribution in Quality Control

Another application of the normal distribution is the analysis of quality control data. It is common practice to include quality control samples into an analytical run to check that the method is performing to specification. The characteristics of the quality control (QC) material are first determined by replicate analyses of the material (usually twenty) then calculating the mean and standard deviation. The same material is then analyzed in each batch of samples (sometimes more than once). If the method is performing to specification then the results for the QC material should fall within the 95% confidence limits (mean ± 2*s*) most of the time. In fact the distribution of results should belong to the same normal distribution as when the characteristics were originally determined i.e. the results should cluster around the mean.

A useful way of plotting the data is to turn the histogram on it's side and construct a *y* axis with horizontal lines representing the mean, 1*s*, -1*s*, 2*s* etc. The results are plotted along the *x* axis and should fall on either side of the mean with equal frequency, most results will fall within the mean ±*s*, fewer results between the *s* and 2*s* limits, very few between 2*s* and 3*s* with only very occasional results outside the 3*s* limits. This is known as a *Levy-Jennings Chart* and an example is shown in Fig 10.4.

The only limitation of using 95% confidence limits (i.e. mean ± 2*s*) as the only criterion is that by definition 1 in 20 analytical runs will be rejected. For a multi-channel analyzer measuring 20 analytes this means that on average one channel will be rejected every run. In other words this criterion is too sensitive. Westgard has devised a set of criteria, the *Westgard Rules*, to improve the power of QC data to detect "real" errors without an unacceptably high lever of "false rejections". These rules are based upon the fact that the probability of obtaining two consecutive results outside the 95% confidence limits is the product of the individual probabilities that one result is outside these limits and is considerably lower and so increases the likelihood that the method is out of control. This idea is extended to four results being between the *s* and 2*s* limits and ten results being one side of the mean etc.

Question Q 10(3)

Calculate the probability of obtaining:
a) a QC result outside the mean ± 3*s* range; b) two QC results between the mean 2*s* and mean 3*s* limits, and c) four results between the mean – *s* and mean – 2*s* limits.

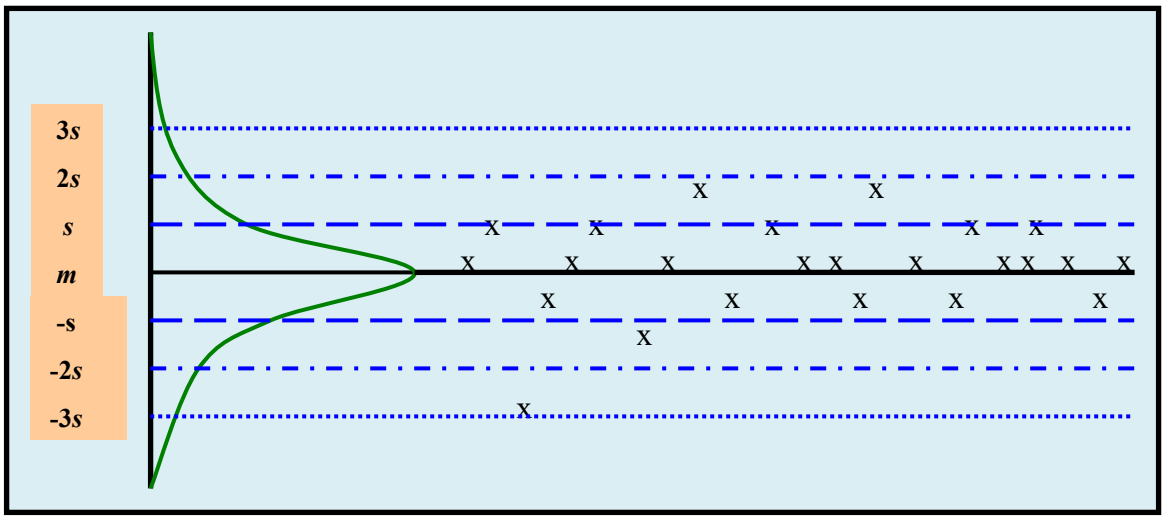


Figure 10.4 Levy-Jennings Quality Control Chart showing 24 sequential results for the same control sample

Answer Q 10(3)

a) The probability of obtaining a result outside the mean ± 3*s* range is the *P* value corresponding to a *z* score of 3. An exact value for z = 3 is not given in Fig 10.4 but the nearest value (*z* = 3.09) can be used as an approximation and corresponds to a *P* value of 0.002. The chance of obtaining a result outside the mean ± 3*s* range is therefore 1 in 1/0.002 which is 1 in 500. This value will occur so infrequently that the method is almost certainly out of control, and is known as the Westgard 1_{3s} rule.

a) The probability of obtaining a result outside the mean ± 2*s* range can again be obtained by looking up the *P* value corresponding to *z* = 2 in Fig 10.4. Again the exact value for *z* = 2 is not given but the nearest (*z* = 1.96) is a good approximation and corresponds to a *P* value of 0.05. There is a slight complication here in that we really wish to know the *P* value when *z* is between ±2 and ±3, not simply when it is greater than ±2. To allow for this all we need to do is to subtract the *P* value for *z* = 3 (P_3) from the *P* value for *z* = 2 (P_2) to give the probability of obtaining a value between 2 and 3 standard deviations from the mean ($P_{2\text{-}3}$):

$$P_{2\text{-}3} = P_2 - P_3 = 0.05 - 0.002 = \mathbf{0.048}$$

Furthermore, we need the probability that the result is between the mean +2*s* and mean +3*s*, not simply between the mean ± 2*s* and mean ± 3s limits on either side of the mean. Therefore the probability of obtaining a result between the mean + 2*s* and mean +3*s* limits is one half of 0.048 i.e. 0.024. The likelihood of obtaining two results between the mean + 2s and mean + 3s limits is the product of the probability of obtaining each result between these limits:

$$0.024 \times 0.024 = \mathbf{0.000576}$$

which is the same as 1 in 1/0.000576 = **1 in 1736**

This is similar to flipping a coin twice. The chance of heads the first time is 0.5 the second time 0.5. Therefore the chance of obtaining heads on both occasions is 0.5 x 0.5 = 0.25. Another way of looking at this is that there are four possible results, heads and tails, heads and heads, tails and heads and tails and tails: there are 4 equally likely results but only one of these is heads on both tosses, so that the probability is 1 in 4 or 0.25).

b) As shown above, the probability (P_{2s}) of obtaining a result outside the mean ± 2*s* limits is 0.05. From Fig 10.4 it can be seen that the chance (P_s) of obtaining a result outside the mean ± *s* range is 0.33. Therefore the probability ($P_{2s\text{-}s}$) of obtaining a result between the ± s and ± 2s limits is given by:

$$P_{2s\text{-}s} = P_s - P_{2s} = 0.33 - 0.05 = 0.28$$

The probability of obtaining a result between the mean –*s* and mean –2*s* is one half of this i.e. 0.14. The likelihood of obtaining four results within this range is obtained by multiplying this probability by itself four times:

$$0.14 \times 0.14 \times 0.14 \times 0.14 = \mathbf{0.00038} \text{ (2 sig figs)}$$

This calculation can be simplified as follows:

$$0.14 \times 0.14 \times 0.14 \times 0.14 = 0.14^4 = \text{antilog}_{10}(4 \times \log_{10} 0.14)$$

$$= \text{antilog}_{10}(4 \times (-0.854)) = \text{antilog}_{10}(-3.42) = 0.00038$$

A probability of 0.00038 is a chance of 1 in 1/0.00038 = **1 in 2632**

Variances are additive

One of the reasons why statisticians prefer to work with squares of deviations rather than directly with deviations is that the former are additive. This is particularly true of the variance i.e. the standard deviation squared (s^2) which is calculated directly from the sum of squares of the deviations about the mean ($\Sigma(x - m)^2$). If there are two independent sources of variation (with variances s_1^2 and s_2^2) contributing to a measurement, then the total variation (with variance s_{total}^2) is described by:

$$s_{total}^2 = s_1^2 + s_2^2 \quad \text{..... Eq. 10.8}$$

It is important to note that it is the variances only (s^2) which are additive, not the standard deviations (s). If we wish to calculate the combined standard deviation from the individual standard deviations then these are first squared, added together then the square root taken of the product:

$$s_{total} = \sqrt{(s_1^2 + s_2^2)} \quad \text{..... Eq. 10.9}$$

Similarly coefficients of variation (cvs) are not additive unless they are first squared:

$$cv_{total} = \sqrt{(cv_1^2 + cv_2^2)} \quad \text{..... Eq. 10.10}$$

One of the most useful applications of this concept in clinical biochemistry is in examining the effects of population variation and analytical impression on laboratory data. If we were to measure the concentration of an analyte, say creatinine, in sera collected from a large number of individuals then it would be possible to calculate the mean, variance etc for the population (s_{total}^2). However, if we were to take one of these samples and analyze it a large number of times, then for that sample we derive both a mean and a variance due to the analytical imprecision of the method ($s_{analytical}^2$).

Clearly the analytical imprecision is contributing to the overall variation in the patient results so that the true population variation, that is the biological variation ($s_{biological}^2$) is much lower. Since the component variances are additive, and assuming that the analytical variance is the same at all analyte concentrations encountered in the population, we can write:

$$s_{total} = \sqrt{(s_{biological}^2 + s_{analytical}^2)} \quad \text{..... Eq. 10.11}$$

The biological variation can be calculated by rearranging this expression:

$$s_{biological} = \sqrt{(s_{total}^2 - s_{analytical}^2)} \quad \text{..... Eq. 10.12}$$

The biological variation can be further subdivided into its components e.g. *intra*-individual and *inter*-individual variations and equation Eq. 10.11 re-written:

$$s_{total} = \sqrt{(s_{intra\text{-}individual}^2 + s_{inter\text{-}individual} + s_{analytical}^2)} \quad \text{..... Eq. 10.13}$$

Another application of this principle is the analysis of the impression of the steps in an analytical process on the total performance. For example, if a method involved pipetting 7 mL of reagent, then this could be achieved by either pipetting 5 mL and 2 mL from separate bulb pipettes or by pipetting 7 mL from a graduated pipette. The combined error from using separate bulb pipettes could be calculated from $s_{total} = \sqrt{(s_{5mL}^2 + s_{2mL}^2)}$ and compared with the variance obtained from using the 7 mL graduated pipette.

Question Q 10(4)

A laboratory using a method with an analytical coefficient of variation of 5% at a concentration of 100 mmol/L for a serum constituent examined samples from a healthy population and found a Gaussian distribution with 95% reference range of 74-126 mmol/L. If the method coefficient of variation had been 22%, what reference range would the laboratory have found?

Answer Q10(4)

The first step is to calculate the overall standard deviation (s_{Total}) and mean (m) from the reference range obtained with the original method (with analytical cv of 5%).

The 95% reference limits incorporates the mean ± 2s, i.e. spans 4s units

$$\text{Therefore } s_{Total} = \frac{\text{Range of 95\% limits}}{4} = \frac{126 - 74}{4} = \frac{52}{4} = 13 \text{ mmol/L}$$

The mean is given by the lower reference limit + 2s.

Therefore, mean (m) = 74 + (2 x 13) = 74 + 26 = 100 mmol/L

Convert the original analytical cv (5%) to its standard deviation ($s_{Analytical}$):

$$\text{Coefficient of variation } (cv\%) = \frac{s \times 100}{m}$$

$$\text{Rearranging, } \quad s = \frac{\text{cv (\%)} \times m}{100}$$

$$\text{Therefore analytical } s \ (s_{Analytical}) = \frac{5 \times 100}{100} = 5 \text{ mmol/L}$$

The measured variation will reflect both the biological and analytical variations. Since it is variances and not standard deviations which are additive, then the square of the total standard deviation (s_{Total}) is equal to the sum of the squares of both the biological ($s_{Biological}$) and analytical ($s_{Analytical}$) standard deviations.:

$$(s_{Total})^2 = (s_{Biological})^2 + (s_{Analytical})^2$$

Substitute s_{Total} = 13 mmol/L and $s_{Analytical}$ = 5 and solve for $s_{Biological}$

$$13^2 = (s_{Biological})^2 + 5^2$$

$$(s_{Biological})^2 = 13^2 - 5^2$$

$$= 169 - 25 = 144$$

$$s_{Biological} = \sqrt{144} = 12 \text{ mmol/L}$$

If new analytical *cv* is 22%, then $s_{Analytical} = \frac{22 \times 100}{100} = 22$ mmol/L

and the new s_{Total} can be calculated (assuming biological variation remains unchanged):

$$s_{Total}^2 = 12^2 + 22^2 = 144 + 484 = 628$$

$$s_{Total} = \sqrt{628} = 25 \text{ mmol/L (to 2 sig figs)}$$

Lower limit of reference range $= m - 2s_{Total} = 100 - (2 \times 25)$

$= 100 - 50 = 50$ mmol/L

Upper limit of reference range $= m + 2s_{Total} = 100 + (2 \times 25)$

$= 100 + 50 = 150$ mmol/L

New reference range = 50 - 150 mmol/L

When is a change in a test result significant?

When monitoring patients it is helpful to know by how much the concentration of any given analyte has to change before the change is clinically significant. To answer this question allowance has to made for the effects of both analytical and within-individual variation. One way to address this problem would be to analyse each sample a number of times and compare their means using a suitable statistical test. In day-to-day practice we do not have this luxury, only single measurements. However, all laboratories should have some idea of the total variability for each of their analytes (which includes both analytical and within-individual variation).

Suppose the result for an analysis is x_1 on the first occasion and x_2 on the second occasion, then we could calculate the difference $(x_1 - x_2)$. We could treat the value for $(x_1 - x_2)$ as a variable which forms a normal Guassian distribution. In other words, if the analysis of the two samples was repeated a large number of times then a histogram could be constructed with frequency plotted against $(x_1 - x_2)$. If the two results (x_1 and x_2 are not significantly different) then their difference $(x_1 - x_2)$ should be within the 95% confidence limits of a distribution with mean of zero and their combined standard deviation ($s_{1,2}$). In other words the difference in results $(x_1 - x_2)$ can be normalized to give a value for z if it is divided by s_{12}:

$$z = \frac{(x_1 - x_2) - 0}{s_{1,2}}$$

The distribution would have a mean of zero with a standard deviation of 1. If there was no significant difference between the results obtained on the two occasions then the peak of the histogram (i.e. the mean value for $(x_1 - x_2)$ would be zero). For a value to be significantly different from the mean (in this case zero) at a probability level of 5% (i.e. $P = 0.05$) the value for z would need to be 1.96. Therefore for the difference $(x_1 - x_2)$ to be significantly different from zero we substitute $z = 1.96$ into the above expression:

$$\mathbf{1.96 = \frac{(x_1 - x_2)}{s_{1,2}}} \quad \text{..... Eq. 10.14}$$

The value for s is actually the combined standard deviations for the two measurements (s_1 and s_2). As seen in the previous section, the combined value for s when two results are combined (i.e. added or subtracted) is the square root of the sum of their squares:

$$s_{1,2} = \sqrt{(s_1^2 + s_2^2)}$$

However, if the value for s is the same at both concentrations then this expression simplifies to:

$$s_{1,2} = \sqrt{(2s^2)} = \sqrt{2} \times s = 1.414\, s$$

Substituting $s_{1,2} = 1.414\, s$ into Eq 10.14 gives:

$$1.96 = \frac{(x_1 - x_2)}{1.414\, s}$$

Which can be rearranged and simplified to:

$$\mathbf{(x_1 - x_2) = 2.8\, s} \quad \text{..... Eq. 10.15}$$

Therefore for a change in a result to be significant at the 5% level of probability the two results must differ by at least $2.8s$. It is important that the value for s is the same at both concentrations.

Question Q 10(5)

While trying to follow the National Service Framework guidelines for coronary heart disease a doctor prescribed a statin to lower the cholesterol of a patient with coronary heart disease. The patient's original cholesterol level was 5.8 mmol/L and at the next visit the doctor was delighted to find that it was just below the target level of 5.0 mmol/L at 4.9 mmol/L and discharged the patient. The patient, a statistician, was less sure the treatment had been responsible. Given that the physiological coefficient of variation for cholesterol is 6% and the analytical coefficient of variation is 3%, calculate the least significant change (at $p<0.05$) in cholesterol as a percentage at his original level, and determine whether the second measurement was significantly different from the first.

Answer Q 10(5)

The total *CV* is the square root of the sum of the squares of the physiological and analytical *CV*s:

$$\text{Total } CV(\%) = \sqrt{\text{Analytical } \%CV^2 + \text{Physiological } \%CV^2}$$

$$= \sqrt{(3^2 + 6^2)} = \sqrt{(9 + 36)} = \sqrt{45} = 6.7\%$$

Next calculate the standard deviation (*s)*:

$$CV(\%) = \frac{s \times 100}{\text{Mean}} \quad \text{therefore } s = \frac{CV(\%) \times \text{mean}}{100}$$

Substitute *CV* = 6.7% and the original level (5.8 mmol/L) as the mean:

$$s = \frac{6.7 \times 5.8}{100} = 0.389 \text{ mmol/L}$$

For two results to be significantly different (at $p < 0.05$) they have to be at least 2.8 standard deviations apart (2.8*s*).

Therefore the least significant change is $2.8 \times 0.389 = 1.09$ mmol/L

Which expressed as a percentage of the original measurement is $\frac{1.09 \times 100}{5.8} = 18.8\%$

Next calculate the difference between the first and second measurement as a percentage of the first measurement:

$$\frac{(5.8 - 4.9) \times 100}{5.8} = 15.5\%$$

which is less than 18.8% so that the change in cholesterol is **not** statistically significant.

ADDITIONAL QUESTIONS

1. The following results were obtained for a QC sample:

   Total protein (g/L): 70, 68, 71, 65, 68, 70, 73, 69, 75, 74, 69, 71

   Calculate the mean, variance, standard deviation, coefficient of variation and 95 per cent confidence limits.

2. Serum thyroxine was measured in 10,000 healthy male adults. Assuming a Gaussian distribution the normal range was calculated to be 50-150 nmol/L. How many results are expected to be above 165 nmol/L?

3. Calculate the least significant difference for a change in cholesterol if the intra-individual coefficient of variation for cholesterol is 4.7% and the analytical coefficient of variation, 2.4%. A patient was changed from Atorvastatin 80 mg to Rosuvastatin 40 mg and the total cholesterol fell from 6.9 to 5.9 mmol/L. Calculate the percentage change in cholesterol and state whether this is significant.

4. Your on-call laboratory service uses 30 different methods, each of which has a 1% probability of failing QC criteria during the course of a night. Assuming that QC of any method is independent of that of the other methods, what is the probability that on any one night all methods will pass the QC criteria?

5. You attempt to derive a reference range for TSH for an ethnic minority population. The first 10 samples give the following results:

Result	n
Between 0.5 and 1.49	5
Between 1.5 and 2.49	3
Between 2.5 and 3.49	0
Between 3.5 and 4.49	1
Between 4.5 and 5.49	1

On the basis of these results, what range of TSH values would encompass 95% of the ethnic minority population?

6. You are required to pipette a 9ml volume and have available a 10 ml graduated pipette which has a 2% CV associated with it's delivery volume and 5 and 2 ml volumetric pipettes each of which has a 1%CV associated with their delivery volumes. What is the error of pipetting a 9 mL volume, expressed as plus/minus mL volume?

a) using the graduated pipette
b) using the volumetric pipettes

7. It has been suggested that a proposed analytical goal for an analyte is that the between batch analytical coefficient of variation should not exceed one half of the "true biological" inter-individual coefficient of variation. Calculate the percentage "expansion" of the measured reference range over the true biological reference range when this analytical goal is exactly met.

Chapter 11

Analysis of means and variances

Comparison of means (the *z*- and *t*-test)

Sometimes we do not want to ask the question "is a single result significantly different from a given population" but "is the mean of a set of results significantly different from the mean of another set". The two problems are approached in a similar manner but difficulties arise because the value of the mean is influenced by the number of results used in its calculation.

a) The standard error

Consider the data in Fig 10.1 for creatinine results obtained with sera from sixty normal individuals (mean = 76 μmol/L; standard deviation = 17 μmol/L). Suppose we were to take two of these results (n=2) *at random* and calculate their mean, then repeat this process a large number of times. The results for the means could then be plotted in the form of a histogram similar to the individual results in Fig 10.1. Such a histogram is called the *sampling distribution of the mean*. The peak value (mean) would be the same but there would be one important difference: the distribution of results would be much narrower and the value of its standard deviation would be lower. The standard deviation of the sampling distribution of the mean is called the *standard error of the mean* (SE_m) to distinguish it from the standard deviation of individual results. This process could be repeated by taking three results (n=3) at random, repeating this process a large number of times and plotting the sampling distribution of the mean in a similar manner. The mean of this new distribution would be unchanged but the distribution of results would be less and their standard deviation (i.e. their standard error) would be lower then when the means of two results were calculated. This process could be repeated with increasing sample size (n) and we would find that as n increases the standard error is reduced. An example is shown in Fig 11.1.

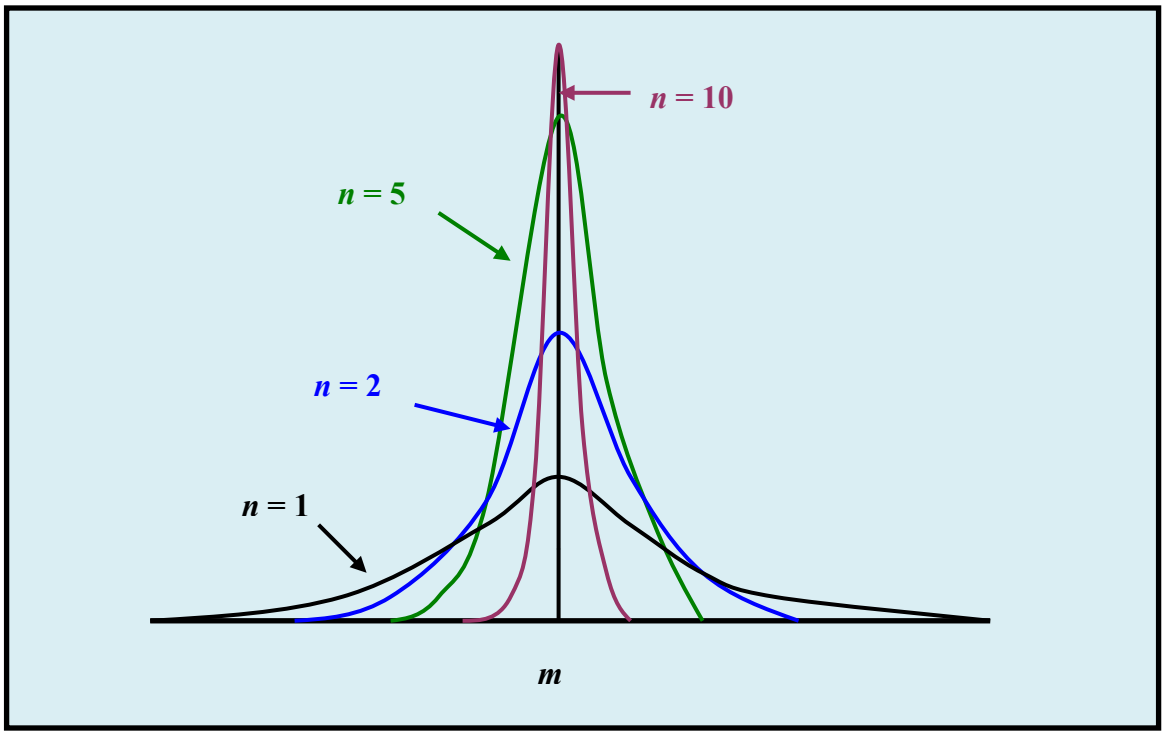


Figure 11.1 Effect of sample size (*n*) on the sampling distribution of the mean

Mathematicians have calculated that the standard error (SE_m) is in fact equal to the standard deviation (s) divided by the *square root* of the number of results (n) used in its calculation:

$$SE_m = \frac{s}{\sqrt{n}} \quad \text{or} \quad \sqrt{\frac{s^2}{n}} \qquad \text{.....Eq. 11.1}$$

Provided the data follow a Guassian distribution the values for the mean (values of m) of samples of size n will be distributed with the peak of the bell shaped curve having the overall true mean (which we shall call μ) and a standard deviation equal to the standard error of the mean (SE_m).

Just as with single measurements, the data can be "normalized" to produce an overall mean of *zero* and a standard error of *one* by calculating the z value:

$$z \text{ (or } t) = \frac{m - \mu}{SE_m} \qquad \text{..... Eq. 11.2}$$

b) Comparing the mean of a sample with a known standard

In order to test whether the mean (m) of a sample size n is significantly different from a hypothetical mean μ of a Guassian distribution first calculate the z (or t) value:

$$z \text{ (or } t) = \frac{m - \mu}{SE_m} = \frac{m - \mu}{s / \sqrt{n}} \quad \text{.....Eq 11.3}$$

then look up the corresponding probability (P value) in tables of z (see Fig 10.4). It is important to note that this approach is only valid if the standard error calculated from the sample (i.e. $s/\sqrt{n}$) is a reasonable estimate of the true standard error of the mean. This in unlikely to be the case unless the value for n is relatively large (greater than 30). William Gossett, who published under the pen-name "*Student*" noted that in small samples, the sample s underestimates the population s. To get around this problem the "t-distribution" was introduced, which is similar to a normal z-distribution in being symmetrical about a mean of zero and is bell-shaped, but differs in that it is flatter (more dispersed) and its dispersion varies according to the size of the sample. The larger the value of n, the more closely a t-distribution resembles a z-distribution. Fortunately statisticians have calculated tables of t for us (portion shown in Fig 11.2) and all that is required is to read off the P-value for the corresponding value of both t and the degrees of freedom (equal to $n - 1$).

Degrees of freedom	Value of P					
	0.10	0.05	0.02	0.01	0.002	0.001
....	.....	.....	.....	.....	.....	.....
....	.....	.....	.....	.....	.....	.....
6	1.943	2.447	3.143	3.707	5.208	5.959
7	1.895	2.365	2.998	3.499	4.785	5.408
8	1.860	2.306	2.896	3.355	4.501	5.041
9	1.833	2.262	2.821	3.250	4.297	4.781
10	1.812	2.228	2.764	3.169	4.144	4.587
.....	.....	.....	.....	.....	.....	.....
30	1.697	2.042	2.457	2.750	3.385	3.646

Figure 11.2 Portion of a table of "Student's" *t*-distribution. e.g. to find the probability (*P*) of obtaining a value for *t* greater than 2.9 for a sample of size 10 (*n*=10), look across the row for 9 degrees of freedom (*n*-1) to find a value for *P*. In this case the next lowest value for *t* (2.821) occurs when *P* is 0.02

b) Comparing the means of two samples

More often we wish to compare two groups of samples (samples 1 and 2), each of which has its own mean (m_1 and m_2), standard deviation (s_1 and s_2) and sample size (n_1 and n_2). Thus the mean of each group of samples has its own standard error ($s_1/\sqrt{n_1}$ and $s_2/\sqrt{n_2}$). In order to calculate a z (or t) value we need to know the combined standard error ($SE_{m1,2}$) i.e. the standard error of the difference between estimates of the two means (m_1-m_2). This combined standard error is calculated in the same was as we calculate the combined standard deviation of biological and analytical variation. Like standard deviations, standard errors are not additive but their squares are (and since we are dealing with squares the signs are always positive):

$$(SE_{m1,2})^2 = SE_{m1}^2 + SE_{m2}^2 = (s_1/\sqrt{n_1})^2 + (s_2/\sqrt{n_2})^2$$

$$= \frac{s_1^2}{(\sqrt{n_1})^2} + \frac{S_2^2}{(\sqrt{n_2})^2} = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$$

Taking square roots gives:

$$\boldsymbol{SE_{m1,2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad \text{..... Eq. 11.4}$$

Therefore dividing the difference between the means by their combined standard error gives the corresponding z (or t value):

$$\boldsymbol{z \text{ (or } t) = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}} \quad \text{..... Eq .11.5}$$

If the values for n_1 and n_2 are greater than 30, then the value for P is obtained from tables of z. If however, the value of n_1 and/or n_2 is less than 30 then the value for P is obtained from tables of t and the degrees of freedom (DF) calculated from the expression:

$$DF = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{[(s_1^2/n_1)^2/(n_1 - 1)] + [(s_2^2/n_2)^2/(n_2 - 1)]} \quad \text{..... Eq. 11.6}$$

For the special case where s_1 and s_2 are equal, simplified versions of these formulae can be used. However, it is first necessary to carry out a *variance ratio test* to see if this is indeed the case so it is probably simplest to stick with only one formula that can be used whatever the relative magnitudes of the individual variances.

Question Q 11(1)

In January a laboratory analysed a quality control sample for sodium 10 times and obtained a mean result of 150 mmol/L with a standard deviation of 4 mmol/L. In February the same sample was analysed 10 times and gave a mean of 154 mmol/L with a standard deviation of 2 mmol/L. Has there been a significant change in performance between January and February?

Answer Q 11(1)

t is calculated using Eq 11.5:

$$t = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where: $m_1 = 150$ mmol/L, $m_2 = 154$ mmol/L
$s_1 = 4$ mmol/L, $s_2 = 2$ mmol/L
$n_1 = 10$, $n_2 = 10$

Therefore $t = \frac{150 - 154}{\sqrt{[(4^2)/10 + (2^2)/10]}}$

$$= \frac{-4}{\sqrt{(1.6 + 0.4)}} = \frac{-4}{\sqrt{2}} = \frac{-4}{1.41} = -2.84$$

Since n_1 and n_2 are small (less than 30), the degrees of freedom (DF) is calculated using Eq. 10.21:

$$DF = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{[(s_1^2/n_1)^2/(n_1 - 1)] + [(s_2^2/n_2)^2/(n_2 - 1)]}$$

Therefore $DF = \frac{(4^2/10 + 2^2/10)}{[(4^2/10)^2/(10-1) + (2^2/10)^2/(10-1)]} = 13$ (2 sig figs)

From tables of t it can be seen that for 13 degrees of freedom, at the 5% level of probability t should be outside the limits -2.16 to +2.16 to be statistically significant. In fact a t value of -2.84 is also significant at the 2% level of probability. Therefore we can conclude that there has been a change in performance between January and February.

Dealing with paired data (the paired *t*-test)

Sometimes individual measurements are made on a series of samples in order to determine if the two measurements differ e.g. when comparing two analytical methods a series of specimens with different values may be analysed by two different methods. Although the two sets of data could be compared using a standard *t*-test this would be a very inefficient way of determining if the two methods of measurement yield different results. This is because the variation in results due to the specimens having different analyte concentrations far outweighs the smaller difference in the two sets of results due to the methods used. A better way to answer this question would be to carry out a paired *t*-test.

For each pair of results, one value is subtracted from the other to give the difference (d). For n pairs of results there will therefore be n values for d. We can calculate the mean (m_d) and standard deviation (s_d) for these values of d. If there is no difference between the two sets of results then the average value of d would be zero. To test whether there is a statistically significant difference of m_d from zero we calculate t in which we are comparing the value for m_d with zero assuming that values for m_d are normally distributed with a standard error of $s_d/\sqrt{n}$:

$$t = \frac{m_d}{s_d/\sqrt{n}} \quad \text{..... Eq 11.7}$$

Where, $m_d = \sum d / n$ and, $s_d = \sqrt{[\sum (d - m_d)^2 / (n-)]}$

Question Q 11(2)

It is suspected that the glucose results obtained with near patient testing (NPT) device on the ward are positively biased. One of the investigations into the problem involves analyzing a series of blood specimens on both the NPT device (A) and an analyzer in the laboratory which measures whole blood glucose (B), with the following results:

Specimen No	1	2	3	4	5	6	7	8	9	10
Glucose (mmol/L) NPT (A)	4.5	6.8	3.2	5.8	8.9	9.5	4.8	7.3	5.1	7.8
Glucose (mmol/L) lab (B)	4.2	7.0	2.8	5.6	8.7	9.7	4.9	6.8	4.6	7.7

Do these results support the suspicion of bias?

Answer Q 11(2)

The variabilities of the results in groups A and B are due to differing glucose concentrations in the specimens **and** to the analytical variation between the instruments. Therefore a standard *t*-test comparing the means of both sets of results would be inappropriate for comparing the analytical performance of method B with method A. As the data are paired, i.e. the same samples were assayed by both instruments, a paired *t*-test can be used.

Calculate the difference (d) between each pair of results: d = A - B

A	B	d	d^2	$d - m_d$	$(d - m_d)^2$
4.5	4.2	0.3	0.09	0.13	0.017
6.8	7.0	-0.2	0.04	-0.37	0.137
3.2	2.8	0.4	0.16	0.23	0.053
5.8	5.6`	` 0.2	0.04	0.03	0.001
8.9	8.7	0.2	0.04	0.03	0.001
9.5	9.7	-0.2	0.04	-0.37	0.137
4.8	4.9	-0.1	0.01	-0.27	0.073
7.3	6.8	0.5	0.25	0.33	0.109
5.1	4.6	0.5	0.25	0.33	0.109
7.8	7.7	0.1	0.01	-0.07	0.005
		$\sum d$ = 1.70	$\sum d^2$ = 0.93		$\sum(d - m_d)^2$ = 0.642

If there is no bias then the differences between each pair of results (d) would be very small and the average would be very close to zero. A paired *t*-test is used to compare the mean difference (i.e. the mean of d) with a hypothetical value of zero taking into account the standard error of the values of d. The mean and standard error of the difference (between values of d) is calculated in the usual way:

$$m_d = \frac{\sum d}{n} = \frac{1.70}{10} = 0.17 \text{ mmol/L}$$

$$s_d = \sqrt{[\sum(d - m_d)^2 / (n\text{-}1)]}$$

$$= \sqrt{(0.642 / (10 - 1)} = \sqrt{0.0713}$$

$$= 0.27 \text{ mmol/L (2 sig figs)}$$

Alternatively, $s_d = \sqrt{[(\sum d^2 - (\sum d)^2/n) / (n-1)]}$

$= \sqrt{[(0.93 - 1.70^2/10) / (10 - 1)]}$

$= \sqrt{[(0.93 - 0.289) / 9]}$

$= \sqrt{0.0712} = 0.27$ mmol/L (2 sig figs)

Next calculate *t*:

$$t = \frac{m_d}{s_d/\sqrt{n}}$$

$$t = \frac{0.17}{0.27/\sqrt{10}} = \frac{0.17 \times \sqrt{10}}{0.27} = \frac{0.17 \times 3.16}{0.27} = \mathbf{1.99}$$

From tables of *t*, for 9 (i.e. *n*-1) degrees of freedom the probability of obtaining a *t* value of 1.99 is greater than 0.05. Therefore, the mean difference (0.17) is NOT significantly different to zero at the 5 per cent level of probability so the data does NOT demonstrate any bias between the two methods.

Variance ratio (the *F*-test)

Sometimes it is not the means that wish to compare but the variation in results. For example, in question *Q 11(1)* a quality control serum was analysed 10 times in January giving a standard deviation (*s*) of 4 mmol/L, and when analysed 10 times again in February the standard deviation was 2 mmol/L. The variance ratio test (*F*-test) can be used to determine whether or not there has been a significant change in precision. The variance ratio is simply the highest variance (s_1) divided by the lowest variance (s_2):

$$\boldsymbol{F = \frac{s_1^2}{s_2^2}} \quad \textbf{..... Eq. 11.8}$$

where s_1 is greater than s_2.

Therefore $F = \frac{4^2}{2^2} = \frac{16}{4} = 4.0$

The next step is to look up the probability of obtaining this value for F from tables of F. Unlike tables of t, the columns refer to the degrees of freedom of s_1 and the rows to the degrees of freedom of s_2. Therefore, there is a separate table for each level of probability (typically 5 and 1%). The degrees of freedom are n –1 for each variance.

DF_2 \ DF_1	....	7	8	9	10	12	15	....
....	....	....	....	....	....	....	....	....
7	....	5.59	4.74	4.35	4.12	3.97	3.87	....
8	....	5.32	4.46	4.07	3.84	3.69	3.58	....
9	....	5.12	4.26	3.86	3.63	3.48	3.37	....
10	....	4.96	4.10	3.71	3.48	3.33	3.22	....
11	....	4.84	3.98	3.59	3.36	3.20	3.09	....
12	....	4.75	3.89	3.49	6.26	3.11	3.00	....
....	....	....	....	....	....	....	....	....

Figure 11.3 Portion of a table of variance ratio ($F = s_1^2/s_2^2$) values, for which degrees of freedom are DF_1 and DF_2 for s_1 and s_2 respectively (where $s_1 > s_2$), corresponding to $P = 0.025$ (2.5%)

Therefore, when $DF_1 = 9$ and $DF_2 = 9$ (since $DF = n - 1 = 10 - 1 = 9$), the probability that the observed value for F is 3.86 is exactly 0.05 (or 5%). Since the F value of 4 is greater than this the probability that it occurred by pure chance is less than 0.05 so that there has been a significant change in precision.

Analysis of variance (ANOVA)

An alternative to comparing means of two sets of data is to analyse the variability that exists within the data. Consider the following two sets of data obtained by analyzing the same QC sample five times in each of two analytical runs (A and *B*):

	A	***B***	
	6.5	*6.0*	
	6.2	*5.8*	
	6.8	*5.4*	
	5.8	*5.6*	
	6.3	*5.9*	
Mean:	6.34	*5.74*	Overall mean = 6.04

Normally the data would be analysed by calculating the mean and variances of data sets A and *B* separately then applying a *t*-test. ANOVA involves calculating the variance of the combined data, but in three separate ways:

Between groups variance The variation between results within each group is eliminated by substituting each group mean for the individual results. In this example group A would consist of 5 results each of 6.34 and group *B* of 5 results each of *5.74* with an overall mean of 6.04.

$$\text{Between groups variance} = [\,(6.34 - 6.04)^2 + (6.34 - 6.04)^2 + (6.34 - 6.04)^2 + (6.34 - 6.04)^2 + (6.34 - 6.04)^2 + (5.74 - 6.04)^2 + (5.74 - 6.04)^2 + (5.74 - 6.04)^2 + (5.74 - 6.04)^2 + (5.74 - 6.04)^2\,] / (n - 1)$$

which can be simplified to:

$$\text{Between groups variance} = [5(6.34 - 6.04)^2 + 5(5.74 - 6.04)^2] / (n-1)$$

Because we have substituted means for individual results, n is only 2 so that n-1 =1.

Therefore between groups variance = 0.45 + 0.45 = **0.90**

Within groups variance This is the variance for the individual results calculated using the appropriate group (A or *B*) mean, but combining both sets of data:

$$\text{Within groups variance} = [\ (6.5 - 6.34)^2 + (6.2 - 6.34)^2 + (6.8 - 6.34)^2 + (5.8 - 6.34)^2 + (6.4 - 6.34)^2 + (6.0 - 5.74)^2 + (5.8 - 5.74)^2 + (5.4 - 5.74)^2 + (5.6 - 5.74)^2 + (5.9 - 5.74)^2\ ] / (n - 2)$$

Although there were 10 results initially, two means were used in the calculation, so that there are $n - 2 = 10 - 2 = 8$ degrees of freedom. Evaluation of the above expression gives a within groups variance of **0.10**.

Sometimes the within groups variance is referred to as the *residual variance*.

Total variance This is the combined variance of the two groups using the overall mean value (in this case 6.04):

$$\text{Total variance} = [(6.5 - 6.04)^2 + (6.2 - 6.04)^2 + (6.8 - 6.04)^2 + (5.8 - 6.04)^2 + (6.4 - 6.04)^2 + (6.0 - 6.04)^2 + (5.8 - 6.04)^2 + (5.4 - 6.04)^2 + (5.6 - 6.04)^2 + (5.9 - 6.04)^2] / (n - 1)$$

There are 10 results initially, one mean was used in the calculation, so that there are $n - 1 = 10 - 1 = 9$ degrees of freedom. Evaluation of the above expression gives a total variance of **0.19**.

The null hypothesis which is used is that if the two sets of data are from the same population then the between groups variance will not be significantly different from the within groups (residual) variance. This can be tested by calculation of the variance ratio:

$$F = \frac{\text{Between groups variance}}{\text{Within groups (residual) variance}} \quad \text{..... Eq. 11.9}$$

In this example, F is 0.90/0.10 = 9.0 with 1 and 8 degrees of freedom. From tables of F at the 5% level of probability the F value would be 5.32. Since the value obtained is higher than this then the two sets of data are significantly different at the 5% level of probability (i.e. P <0.05). ANOVA for two sets of data is rarely used since it is much easier to compare the means directly with a *t*-test. In fact the P value obtained by ANOVA is exactly the same as that obtained with a *t*-test.

The value of ANOVA comes into its own when comparing more than two sets of data. For example, if we had four sets of data we wished to compare, A, B, C and D then one option would be to carry out *t*-test between each possible combination of data. The various combinations are: A-B, A-C, A-D, B-C, B, D and C-D making six *t*-tests in all. Clearly it would be simpler to first carry out an ANOVA to se if a difference exists between *any* of the groups of data. There is another reason for using ANOVA in preference to multiple *t*-tests. If we are looking for a difference which is significant at the 5% level and no significant difference really exists between two sets of data then a significant value for *t* will be obtained on five occasions out of a hundred by chance alone i.e. a false positive rate of 5% will be obtained. If numerous *t*-tests are carried out then the incidence of false positives is even higher.

An underlying assumption when carrying out ANOVA is that the variances of the individual groups are homogeneous i.e. they are not significantly different from each other. This can be confirmed by first carrying out a variance ratio test on the variances of the two groups which have the highest and lowest variance. If the variances are not homogeneous then the problem can often be overcome by first transforming the data.

A difficulty which often arises is that the size of the groups are not equal i.e. some of the data is missing. Computer packages, which are usually used to perform these calculations nowadays, have the facility to correct for missing data.

If ANOVA does not reveal a significant difference between the groups then further statistical analysis is not usually necessary. However, if a difference is demonstrated then ANOVA cannot tell us which group(s) is/are significantly different from the rest. This can only be done by carrying out appropriate *t*-tests – although simple inspection of the data usually suggest which group(s) is/are likely to be different.

It is common practice to first calculate sums of squares rather then variances. This is because sums of squares are additive e.g. the between groups and within groups sums of squares should add up to the total – a useful check on the calculations (a similar check can be applied to degrees of freedom). In fact it is only necessary to calculate two of the sums of squares e.g. the total and within groups, then obtain the between groups sum of squares by subtraction.

1. Let u = number of groups of data and v = the number of data points in each group. Arrange the data (values of x) into a table with u columns and v rows.

2. At the foot of each column enter values for $\sum x$, n, *mean*, $\sum x^2$, $(\sum x)^2/n$ and sum of squares [i.e. $\sum x^2 - (\sum x)^2/n$] *for each group*.

		Group No						
		1	2	3	4	…	u	**SUM**
	1	$x_{1,1}$	$x_{2,1}$	$x_{3,1}$	$x_{4,1}$	….	$x_{u,1}$	
Sample No	2	$x_{1,2}$	$x_{2,2}$	$x_{3,2}$	$x_{4,2}$	….	$x_{u,2}$	
	3	$x_{1,3}$	$x_{2,3}$	$x_{3,3}$	$x_{4,3}$	….	$x_{u,3}$	
	4	$x_{1,4}$	$x_{2,4}$	$x_{3,4}$	$x_{4,4}$	….	$x_{u,4}$	
	…	….	….	….	….	….	….	
	v	$x_{1,v}$	$x_{2,v}$	$x_{3,v}$	$x_{4,v}$	….	$x_{u,v}$	
$\sum x$	…	…	…	…	…	…	…	$\sum . \sum x$
n	…	…	…	…	…	…	…	u x v
Mean	…	…	…	…	…	…	…	
$\sum x^2$	…	…	…	…	…	…	…	$\sum . \sum x^2$
$(\sum x)^2/n$	…	…	…	…	…	…	…	$\sum . (\sum x)^2/n$
$\sum x^2 - (\sum x)^2/n$	…	…	…	…	…	…	…	

3. Calculate the totals, *for each group*, of $\sum x$, $\sum x^2$ and $(\sum x^2)/n$ and enter in a new column to the right. Complete the following table using these values:

Source of Variance	**Sum of squares (*SS*)**	**Degrees of freedom (*DF*)**	**Mean square (Variance,s^2)**
Between groups	$\sum . (\sum x)^2/n \; - \; (\sum . \sum x)^2/uv$	u - 1	*SS*/*DF*
Within groups	$\sum . \sum x^2 \; - \; \sum . (\sum x)^2/n$	$u(v$ -1)	*SS*/*DF*
Total	$\sum . \sum x^2 \; - \; (\sum . \sum x)^2/uv$	(uv) - 1	*SS*/*DF*

4. Divide the between groups mean square by the within groups mean square to obtain F and look up its P value in tables of F with u - 1 and $u(v$ - 1) degrees of freedom.

Figure 11.4 Simplified procedure for performing one-way ANOVA

The example discussed above in which the data was divided into several sets is known as one-way analysis of variance. This simple concept can be extended to two-way, three-way etc analysis of variance. For example if we were conducting a study to compare various drug treatments with a group of patients receiving each treatment then we would need to carry out a one-way ANOVA. If however, the patients were divided into males and females then the creation of these extra groups would require a two-way ANOVA. Details of these techniques can be found in statistics textbooks, but nowadays computer packages are usually used to perform these calculations.

Question Q11(3)

A laboratory was attempting to optimize a new alkaline phosphatase assay. Four different buffers (A, B, C and D) were each used to assay the same serum sample ten times with the following results (expressed as IU/L):

	Buffer		
A	**B**	**C**	**D**
175	170	175	168
160	162	168	158
162	180	198	183
189	165	174	174
177	165	178	178
165	158	182	162
171	164	184	176
190	191	178	193
162	176	201	168
170	168	194	175

Is there a significant difference in the measured activity between any of the four buffers?

Answer Q11(3)

There are two methods available for calculation of the between groups, within groups and total sums of squares:

Method 1 Direct calculation from the means

First calculate the mean of each column and the overall mean:

$\text{Mean}_A = (175 + 160 + 162 + \ldots\ldots..162 + 170)/10 = 172.1$

$\text{Mean}_B = (170 + 162 + 180 + \ldots\ldots..176 + 168)/10 = 169.9$

$\text{Mean}_C = (175 + 168 + 198 + \ldots\ldots..201 + 194)/10 = 183.2$

$\text{Mean}_C = (168 + 158 + 183 + \ldots\ldots..168 + 175)/10 = 173.5$

$\text{Mean}_{(A..D)} = (172.1 + 169.9 + 183.2 + 173.5)/4 = 174.7$

The between groups sum of squares is calculated from the group means and overall means:

$$\text{Between groups sum of squares} = 10\,[(172.1 - 174.7) + (169.9 - 174.7) + (183.2 - 174.7) + (173.5 - 174.7)]$$

$$= 1035 \quad [\text{degrees of freedom} = 4 - 1 = 3]$$

The within groups sum of squares is calculated from the individual results and the group means:

$$\text{Within-groups sum of squares} = (175 - 172.1) + (160 - 172.1) + \ldots\ldots..(170 - 172.1) + (170 - 169.9) + (162 - 169.9) + \ldots\ldots\ldots(168 - 169.9) + (175 - 183.2) + (168 - 183.2) + \ldots\ldots\ldots(194 - 183.2) + (168 - 173.5) + (158 - 173.5) + \ldots\ldots\ldots(175 - 173.5)$$

$$= 3944 \quad [\text{degrees of freedom} = 4\,(10 - 1) = 36]$$

The total sum of squares is calculated from the individual results and the overall mean:

Total sum of squares = (175 - 174.7) + (160 - 174.7) + ……. (170 - 174.7)

+ (170 - 174.7) + (162 - 174.7) + …… (168 - 174.7)

+ (175 - 174.7) + (168 - 174.7) + …… (194 - 174.7)

+ (168 - 174.7) + (158 - 174.7) + ……. (175 - 174.7)

= 4978 [degrees of freedom = (4 x 10) - 1 = 39]

Method 2. Via individual sums of squares

This is the procedure shown in Fig 11.4:

	Buffer				
	A	**B**	**C**	**D**	
	175	170	175	168	
	160	162	168	158	
	162	180	198	183	
	189	165	174	174	
	177	165	178	178	
	165	158	182	162	
	171	164	184	176	
	190	191	178	193	
	162	176	201	168	
	170	168	194	175	
					TOTALS
$\sum x$	1721	1699	1832	1735	6987
n	10	10	10	10	40
Mean	172.1	169.9	183.2	173.5	698.7
$\sum x^2$	297229	289535	336714	301955	1225433
$(\sum x)^2/n$	296184	288660	335622	301023	1221489
$\sum x^2 - (\sum x)^2/n$	1045	875	1092	933	3944

Number of groups (u) = 4, number in each group (v) = 10.

Between groups sum of squares = $\sum.(\sum x)^2/n - (\sum.\sum x)^2/uv$

= $\underline{1221489} - (6987^2/40)$ = 1035

Within groups sum of squares = $\sum.\sum x^2 - \sum.(\sum x)^2/n$

= 1225433 - 1221489 = 3944

Total sum of squares = $\sum.\sum x^2 - (\sum.\sum x)^2/uv$

= 1225433 - $(6987^2/40)$ = 4979

Whichever method was used to calculate the sum of squares, the procedure to calculate the F value is the same:

Source	Sum of squares	DF	Variance	Variance ratio (F)
Between groups	1035	3	345	3.15
Within groups	3944	36	110	
Total	4979	39	128	

From tables the probability of obtaining an F value of greater than 2.84 (for 3 and 36 degrees of freedom) is 0.05 (5 per cent). Therefore the data are not homogeneous i.e. at least one of the groups of data are significantly different to the rest. ANOVA cannot tell us which group(s) is/are different, t-tests must be performed on paired groups of data.

FURTHER QUESTIONS

1. The following analytical results were obtained on the same QC sample: 109, 91, 105, 112, 90, 115, 89, 113, 93, 94. Calculate the mean, standard deviation and standard error of the mean.

2. Two laboratories measured sodium in the same plasma sample ten times. One laboratory obtained a mean of 145 mmol/L with an SD of 3 mmol; the other obtained a mean of 147 mmol/L with an SD of 2 mmol/L. Do the laboratories differ in their bias or imprecision?

3. Serum thyroxine was measured in 500 healthy adults. Assuming a Gaussian distribution, the normal range was calculated to be 50-150 nmol/L. What is the probability that the mean of a set of 9 results taken at random from this population is greater than 125 nmol/L?

4. It is suspected that an instrument used for near patient measurement of cholesterol is showing positive bias. The following data are the results of paired analyses of samples from ten patients measured on the standard laboratory analyser (A) and the instrument under investigation (B). Assuming that the results from the main analyser are correct, is there any evidence of bias?

A	B
6.8	7.2
4.2	4.5
5.0	4.8
5.6	5.9
8.5	8.7
2.9	2.8
4.8	4.9
7.6	8.1
6.5	6.4
5.0	5.2

5. Four laboratories in a managed network compared the performance of their serum cholesterol assays by measuring the same sample 10 times with the following results:

Lab			
A	B	C	D
7.6	7.5	7.3	7.7
7.3	7.6	7.4	7.8
7.5	7.2	7.7	7.4
7.7	7.5	7.8	7.5
7.5	7.7	7.4	7.2
7.6	7.4	7.2	7.5
7.4	7.8	7.5	7.3
7.8	7.5	7.6	7.6
7.2	7.3	7.5	7.6
7.5	7.4	7.6	7.4

Is there any significant difference in bias for serum cholesterol at this concentration between the four laboratories?

Chapter 12

Correlation and regression

During graphical analysis of laboratory data the question often arises as to whether or not there is a valid relationship between variables and, if there is, where the line of best fit should be drawn? The statistical techniques of correlation and regression seek to answer these questions. Provided there is a linear relationship between two variables then the best fit will be a straight line. Often the data is best described by a curve in which case more advanced techniques are needed which are beyond the scope of this book. Fortunately it is often possible to transform non-linear data (e.g. by taking logarithms or reciprocals) to a reasonably straight line which can be fitted by linear techniques.

The graphical presentation of data

It is always a good idea to first plot the data before proceeding to statistical analysis. This is to check that the data is likely to fit a linear equation, to detect outliers and confirm that there is a reasonable spread of data points throughout the range for both sets of values. It is customary to use the *horizontal axis* for the *independent* variable which is often denoted x. The *vertical* axis is usually used for the *dependent* variable which is often denoted y. The independent variable is the one whose value is accurately known and the dependent variable is some *function* of the dependent variable. Sometimes it is not obvious which are the dependent and independent variables.

It is customary to write a linear equation in the form:

$$y = bx + a \quad \text{..... Eq. 12.1}$$

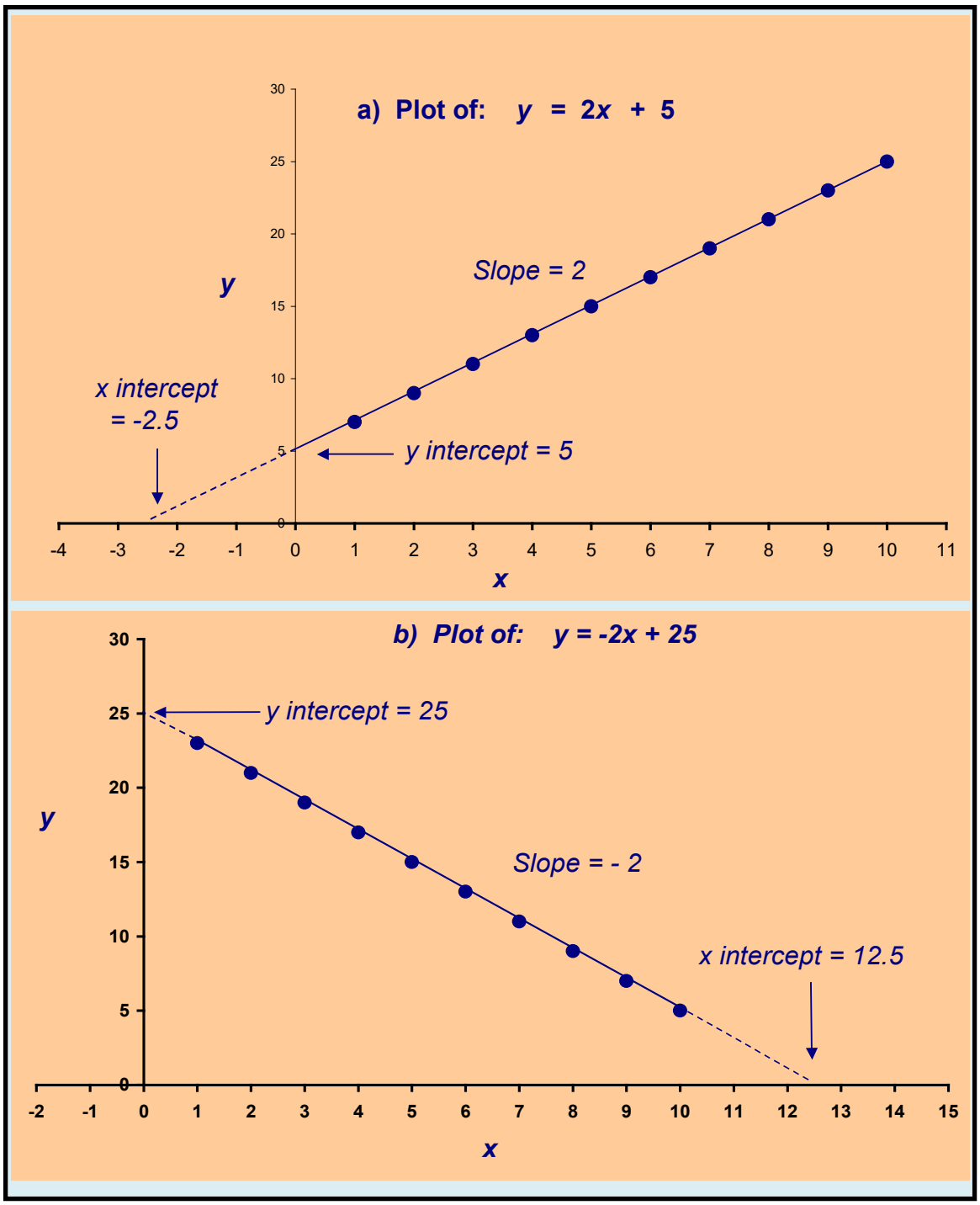


Figure 12.1 **Plots of two equations showing the significance of the slope and intercept**

Where x and y are the independent and dependent variables respectively and a and b are constants. b is the slope of the line i.e. the rate of change of y with x. b can be determined by drawing a right-angled triangle with the line through the data as the hypoteneuse and measuring the vertical and horizontal sides and dividing the former by the latter. Alternatively the angle between the line through the data with the base of the triangle can be measured and its tangent calculated. It is important to take into account the scales for x and y. a is a constant and represents the intercept on the y axis if the line through the data is extrapolated. If a has a value of zero then the line passes through the origin i.e. where the x and y axes intercept.

Figure 12.1 shows a plot of two sets of data of y against x. In both cases the x values are 1, 2, 3,…..10. In (a) the values of y are calculated according to the equation $y = 2x + 5$. Therefore the slope of the line is 2 i.e. when x increases by a value of 1, y increases by a value of 2. The value of the y intercept is 5. This is the value of y when $x = 0$ in which case the expression is $y = (2 \times 0) + 5$ which simplifies to $y = 5$. When $y = 0$ however, the expression becomes $0 = 2x + 5$ which can be rearranged to $2x = -5$ so that $x = -5/2 = -2.5$ which is the intercept on the x axis.

In (b) the values for y are calculated according to the equation $y = -2x + 25$. Therefore the slope of the line is -2 i.e. as x increases by 1, y *decreases* by a value of 2. The value of the y intercept is 25. This is the value of y when $x = 0$ in which case the expression is $y = (-2 \times 0) + 25$ which simplifies to $y = 25$. When $y = 0$ however, the expression becomes $0 = -2x + 25$ which can be rearranged to $2x = 25$ so that $x = 25/2 = 12.5$ which is the intercept on the x axis.

The correlation coefficient (r)

The Pearson *correlation coefficient* (r) is a number between -1 and +1 whose sign is the same as the slope of the best fit line to the data and the magnitude of which is related to the degree of linear association between the two variables. Figure 12.3 shows a plot in which all points fall *exactly* on a line with a *positive* slope so that $r = +1$. Figure 12.4 shows a similar plot in which all the points fall *exactly* on a line with *negative* slope so that $r = -1$. However, the data in Fig 12.5 are scattered widely so that there is no linear relationship between x and y and as a result $r = 0$. In practice r values are usually encountered between 0 and 1 or 0 and -1: the nearer r is to 1 (or -1) the more significant the linear relationship between the two variables.

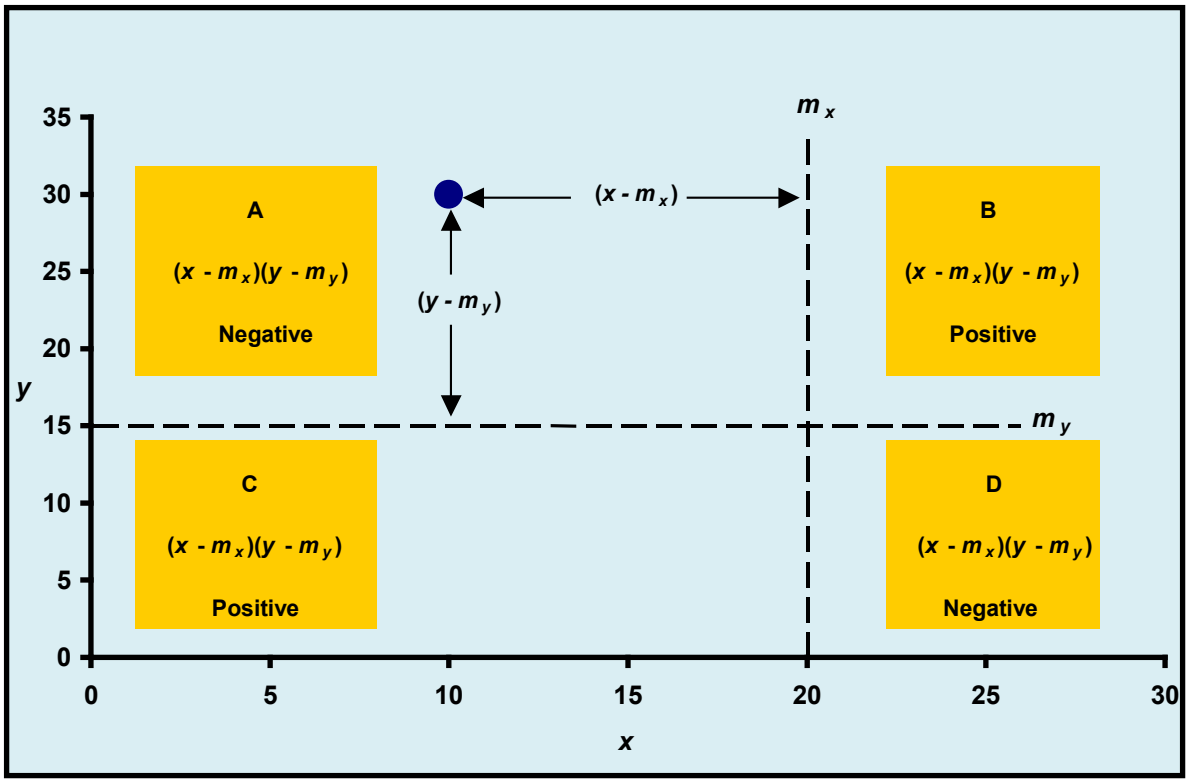


Figure 12.2 **Computation of $(x - m_x)$ and $(y - m_y)$ for a single point (part of a set of data). The intersecting axes representing the means (m_x and m_y) divide the graph area into 4 quadrants in which the product $(x - m_x)(y - m_y)$ is either positive or negative**

The idea behind the correlation coefficient is that it should measure the degree of *association* between x and y. The measure used is the *sum of the products* of the individual deviations of x from its mean and of y from its mean. Figure 12.2 shows the mode of calculation of both deviations for a single point (which is only one of a series of values for x and y). A vertical axis is drawn to represent the mean of the values of x (m_x) and a horizontal axis to represent the mean of all the values of y (m_y). The intersection of these two axes *always* falls on the best straight-line fit to the data. The distance of the observed value of x (10) from the mean of all values of x (20) i.e. $(x - m_x)$, is 10 – 20 = -10. The distance of the observed value of y (30) from the mean of all values of y (15) i.e. $(y - m_y)$ is 30 – 15 = 15. Their product, $(x - m_x)(y - m_y)$ is therefore –10 x 15 = -150.

Note that the intersecting axes of m_x and m_y divide the area of the plot into four quadrants – A, B, C and D. In quadrants A and D the product of the deviations is always negative, whereas in quadrants B and C it is always positive. The product of each pair of deviations may be thought of as a "turning moment" about the intersection point of the m_x and m_y axes. In fact r is often called *Pearson's product-moment coefficient*. If the sum of all the values of $(x - m_x)(y - m_y)$ is divided by the number of data points (or more specifically the number of data points minus one) then the resulting parameter is known as the *covariance* of x and y (c_{xy}):

$$c_{xy} = \frac{\sum(x - m_x)(y - m_y)}{n - 1}$$

Division by the square root of the products of the individual variances of x and y:

$$s_x^2 = \frac{\sum(x - m_x)^2}{n - 1} \quad \text{and} \quad s_y^2 = \frac{\sum(y - m_y)^2}{n - 1}$$

corrects for the total variability in the data and yields an expression for Pearson's correlation coefficient (r):

$$r = \frac{c_{xy}}{\sqrt{s_x^2 . s_y^2}}$$

Substitution of the expressions for c_{xy}, s_x^2 and s_y^2 gives:

$$r = \frac{\sum(x - m_x)(y - m_y)/(n - 1)}{\sqrt{\{[\sum(x - m_x)^2/(n - 1)][\sum(y - m_y)^2/(n - 1)]\}}}$$

Cancelling the $(n - 1)$ values gives the following simpler expression for r:

$$\boldsymbol{r = \frac{\sum(x - m_x)(y - m_y)}{\sqrt{[\sum(x - m_x)^2 . \sum(y - m_y)^2]}}} \quad \text{..... Eq. 12.2}$$

The following version of this expression is often easier to evaluate:

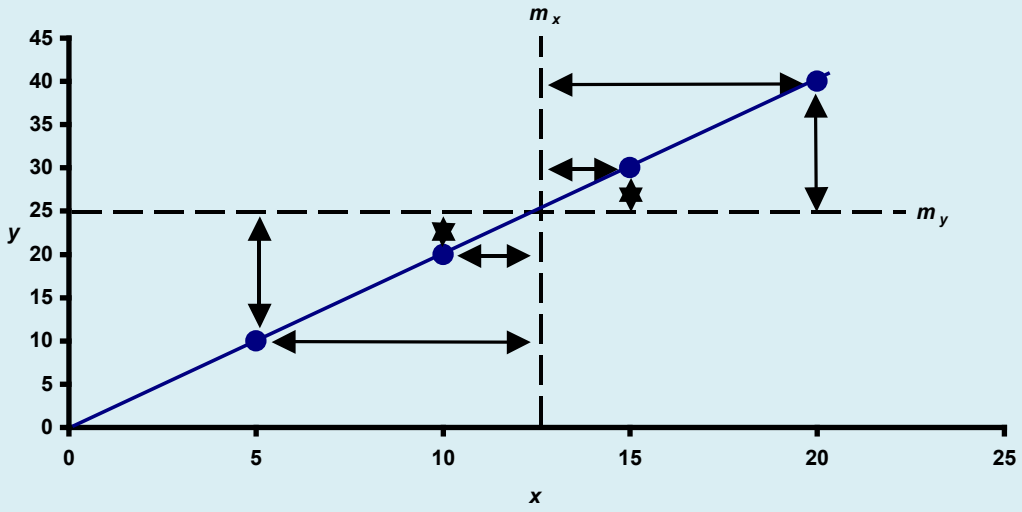
$$r = \frac{\sum xy - \sum x \sum y/n}{\sqrt{\{[\sum x^2 - (\sum x)^2/n][\sum y^2 - (\sum y)^2/n]\}}} \quad \text{..... Eq. 12.3}$$

The way in which the products of the two deviations affect r is illustrated in Figs 12.3 to 12.6. These figures all use the same values for x (i.e. 5, 10, 15 and 20) and y (10, 20, 30 and 40) but differ in the way each value of y is matched with a value of x. Since all the values for x and y are the same it follows that $\sum(x - m_x)^2$ and $\sum(y - m_y)^2$, and therefore $\sqrt{[\sum(x - m_x)^2 \sum(y - m_y)^2]}$ are also identical. However, the value of $\sum(x - m_x)(y - m_y)$ varies according to the pairing of x and y. This is not surprising since the product $(x - m_x)(y - m_y)$ will not only depend upon the value of y but the value of x with which it is associated.

In Fig 12.3 the values of x and y are matched so that all the data points fall on a straight line described by the relationship $y = 2x$. All data points fall into quadrants B and C so that each value of $(x - m_x)(y - m_y)$ and hence r is positive. In fact the values for $\sum(x - m_x)(y - m_y)$ and $\sqrt{[\sum(x - m_x)^2 \sum(y - m_y)^2]}$ are identical (= 250) so that $r = +1$. However, in Fig 12.4 the values for x and y are matched in such a way that the data points all fall on a straight line described by $y = 50 - 2x$. All the points fall into quadrants A and D so that each value of $(x - m_x)(y - m_y)$ and hence r is negative. The value for $\sum(x - m_x)(y - m_y)$ is –250 whereas the value for $\sqrt{[\sum(x - m_x)^2 \sum(y - m_y)^2]}$ is +250 so that their ratio (r) is -1.

In fig 12.5 the values of x and y are matched in such a way that there is no linear relationship between x and y. All four data points fall into different quadrants giving two negative and two positive results for $(x - m_x)(y - m_y)$ which cancel each other exactly. Therefore the value for $\sum(x - m_x)(y - m_y)$, and hence r, is zero.

Figure 12.6 shows the situation more commonly encountered where x and y are matched in such a way that the data can be described by a linear expression (in this case $y = 2x$) but with none of the points falling on the fitted line. All of the points still fall into quadrants B and C but the matching of the values of x and y still results in lower values for $(x - m_x)(y - m_y)$ giving a total of 150. Hence r is 150/250 = 0.6.

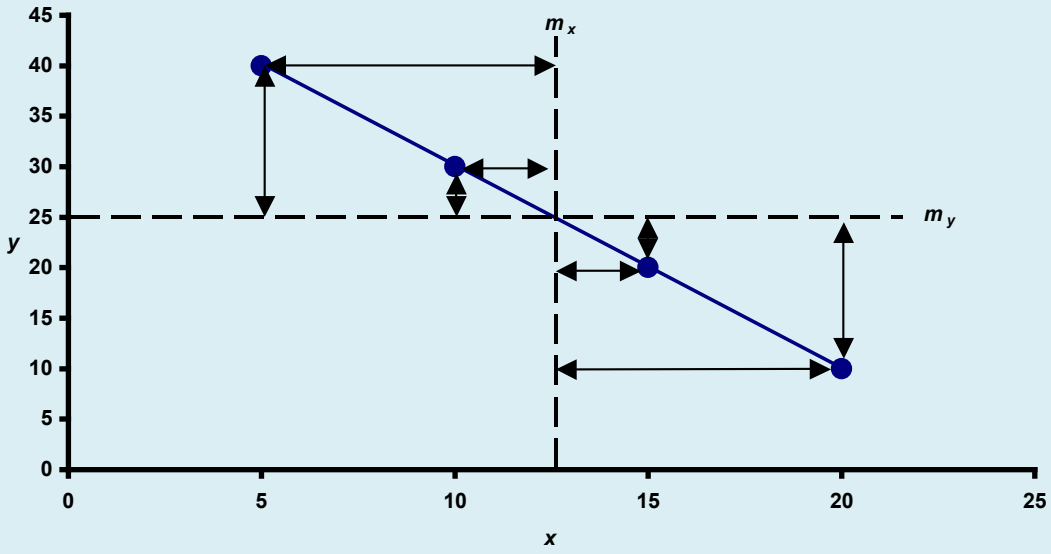


x	y	$(x - m_x)$	$(x - m_x)^2$	$(y - m_y)$	$(y - m_y)^2$	$(x - m_x)(y - m_y)$
5	10	-7.5	56.25	-15	225	112.5
10	20	-2.5	6.25	-5	25	12.5
15	30	2.5	6.25	5	25	12.5
20	40	7.5	56.25	15	225	112.5
$\sum =$ 50	100	0	125	0	500	250

$$m_x = \frac{\sum x}{n} = \frac{50}{4} = 12.5 \; ; \qquad m_y = \frac{\sum y}{n} = \frac{100}{4} = 25$$

$$r = \frac{\sum (x - m_x)(y - m_y)}{\sqrt{[\sum (x - m_x)^2 \sum (y - m_y)^2]}} = \frac{250}{\sqrt{[125 \times 500]}} = \frac{250}{250} = 1$$

Figure 12.3 **A set of four data points which fall exactly on a straight line described by the relationship $y = 2x$ yielding a correlation coefficient (r) of one, showing the method for calculation of deviations and sums of squares and products**

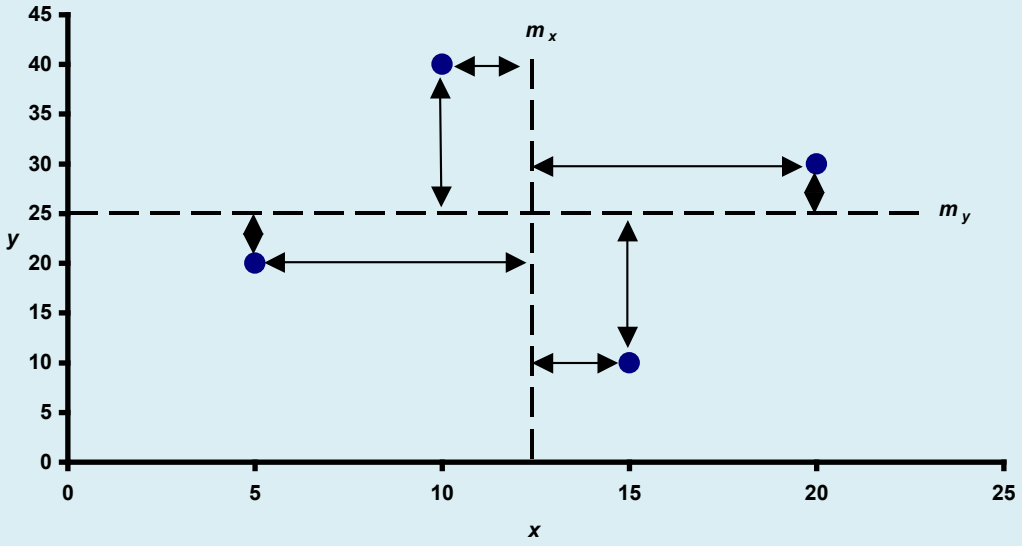


	x	y	$(x - m_x)$	$(x - m_x)^2$	$(y - m_y)$	$(y - m_y)^2$	$(x - m_x)(y - m_y)$
	5	40	-7.5	56.25	15	225	-112.5
	10	30	-2.5	6.25	5	25	-12.5
	15	20	2.5	6.25	-5	25	-12.5
	20	10	7.5	56.25	-15	225	-112.5
$\sum$ =	50	100	0	125	0	500	-250

$$m_x = \frac{\sum x}{n} = \frac{50}{4} = 12.5; \qquad m_y = \frac{\sum y}{n} = \frac{100}{4} = 25$$

$$r = \frac{\sum(x - m_x)(y - m_y)}{\sqrt{[\sum(x - m_x)^2 \sum(y - m_y)^2]}} = \frac{-250}{\sqrt{[125 \times 500]}} = \frac{-250}{250} = -1$$

Figure 12.4 **A set of four data points which fall exactly on a straight line described by the relationship $y = 50 - 2x$ yielding a correlation coefficient (r) of minus one, showing the method for calculation of deviations and sums of squares and products**

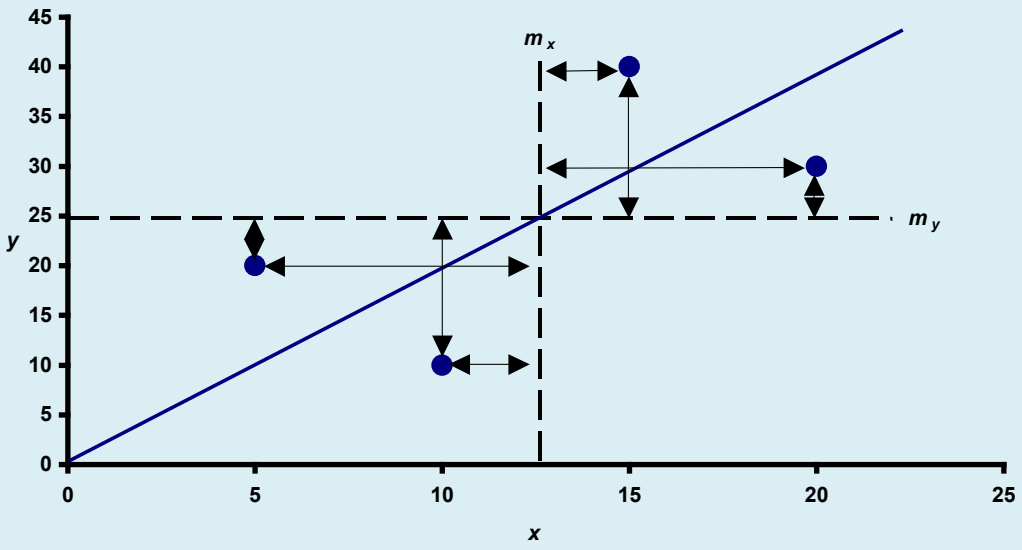


	x	y	$(x - m_x)$	$(x - m_x)^2$	$(y - m_y)$	$(y - m_y)^2$	$(x - m_x)(y - m_y)$
	5	20	-7.5	56.25	-5	25	37.5
	10	40	-2.5	6.25	15	225	-37.5
	15	10	2.5	6.25	-15	225	-37.5
	20	30	7.5	56.25	5	25	37.5
$\sum$ =	50	100	0	125	0	500	0

$$m_x = \frac{\sum x}{n} = \frac{50}{4} = 12.5; \qquad m_y = \frac{\sum y}{n} = \frac{100}{4} = 25$$

$$r = \frac{\sum(x - m_x)(y - m_y)}{\sqrt{[\sum(x - m_x)^2 \sum(y - m_y)^2]}} = \frac{0}{\sqrt{[125 \times 500)}} = \frac{0}{250} = 0$$

Figure 12.5 **The same set of data used in Figs 12.3 and 12.4 but paired in such a way as to yield no correlation (r = 0), showing the method for calculation of deviations and sums of squares and products. Note that each data point falls into a different quadrant**



x	y	$(x - m_x)$	$(x - m_x)^2$	$(y - m_y)$	$(y - m_y)^2$	$(x - m_x)(y - m_y)$
5	20	-7.5	56.25	-5	25	37.5
10	10	-2.5	6.25	-15	225	37.5
15	40	2.5	6.25	15	225	37.5
20	30	7.5	56.25	5	25	37.5
$\sum$ = 50	100	0	125	0	500	150

$$m_x = \frac{\sum x}{n} = \frac{50}{4} = 12.5; \quad m_y = \frac{\sum y}{n} = \frac{100}{4} = 25$$

$$r = \frac{\sum(x - m_x)(y - m_y)}{\sqrt{[\sum(x - m_x)^2 \sum(y - m_y)^2]}} = \frac{150}{\sqrt{[125 \times 500]}} = \frac{150}{250} = 0.6$$

Figure 12.6 **The same set of four data points used in Fig 12.3 to but with the *y* values interchanged to give a poor correlation (r = 0.6) yet still fitting the function $y = 2x$, showing the method for calculation of deviations and sums of squares and products. Note that all four data points still fall into the positive quadrants**

The standard error of r (SE_r) is given by the surprisingly simple expression:

$$SE_r = \sqrt{\frac{1 - r^2}{n - 2}} \quad \text{..... Eq. 12.4}$$

Division of r by SE_r yields an expression for t:

$$t = \frac{r}{SE_r} = \frac{r\sqrt{(n-2)}}{\sqrt{(1 - r^2)}} \quad \text{..... Eq. 12.5}$$

The probability (P) of obtaining this value for t can be obtained from standard tables of t where there are $n - 2$ degrees of freedom. Alternatively tables of r are available from which the value of P can be read directly for any number of degrees of freedom.

A significant correlation only means that there is an association between x and y, it does NOT necessarily mean that a change in x causes a change in y. In other words *correlation does not equal causation.*

Nowadays these calculations are usually performed on a pocket calculator or with a computer statistics package.

Another way of analysing comparison data is to use the analysis of variance approach described in chapter 11. For each pair of values of x and y there is a value for y (which we shall call y_{fit}) which falls exactly on the line of best fit (of course if the x,y data point happens to fall on the line of best fit then $y = y_{fit}$). The sum of squares (SS) for the regression is the sum of the differences of y_{fit} from the horizontal axis described by m_y:

$$SS_{\text{regression}} = \sum(y_{fit} - m_y)^2$$

This tells us how far the predicted values differ from the overall mean (analogous to the between sum of squares used in chapter 11).

The residual sum of squares reflects the difference between the original data and the fitted line:

$$SS_{\text{residual}} = \sum(y - y_{fit})^2$$

These sum of squares can be used to calculated mean squares and hence a value for the variance ratio (F) which can in turn be used to test the hypothesis that the line of best fit is significantly different from the horizontal axis. Alternatively the sum of squares can be used to calculate the *coefficient of determination* (R^2) which is simply the proportion of the total variance described by regression:

$$R^2 = \frac{SS_{\text{regression}}}{SS_{\text{regression}} + SS_{\text{residual}}}$$

Fortunately R^2 turns out mathematically to be the *square* of the correlation coefficient:

$$\boldsymbol{R^2 = r^2 = \frac{[\sum (x - m_x)(y - m_y)]^2}{[\sum (x - m_x)][\sum (y - m_y)]}} \quad \textbf{..... Eq. 12.6}$$

R^2 expresses the proportion of variance of the dependent variable explained by the independent variable. For example, if $R^2 = 0.75$ then 75% of the variation in y is accounted for by the variation in x. As for r a perfect correlation would have a value of 1 (since the square of one is one) and if there is no relationship then R^2 equals zero. Note that as R^2 is the square of the correlation coefficient its value is always positive regardless of whether the slope of the line is positive or negative.

Linear regression

Whilst correlation seeks to establish the degree of relationship between variables, regression analysis attempts to determine the expression which best describes the relationship i.e. the line of best fit. One approach would be to manually draw the line of best fit by eye then measure the slope and intercept of the line to determine the constants in the linear equation. However, this would be subject to operator error and would give no measure of the reliability of the constants obtained.

Since x is the independent variable it is assumed to be without error. Most values for y will not fall on the line of best fit due to inherent imprecision of y. However, the linear relationship between x and y predicts the expected value of y (which we shall call y_{fit}) for any given value of x:

$$y_{fit} = b\,x + a$$

The deviation of the observed value for y from the regression line is $(y - y_{fit})$, which is also known as the *residual* (e), and can be either positive or negative depending upon whether the observed value falls above or below the line (see Fig 12.7):

$$e = y - y_{fit} = y - b\,x - a$$

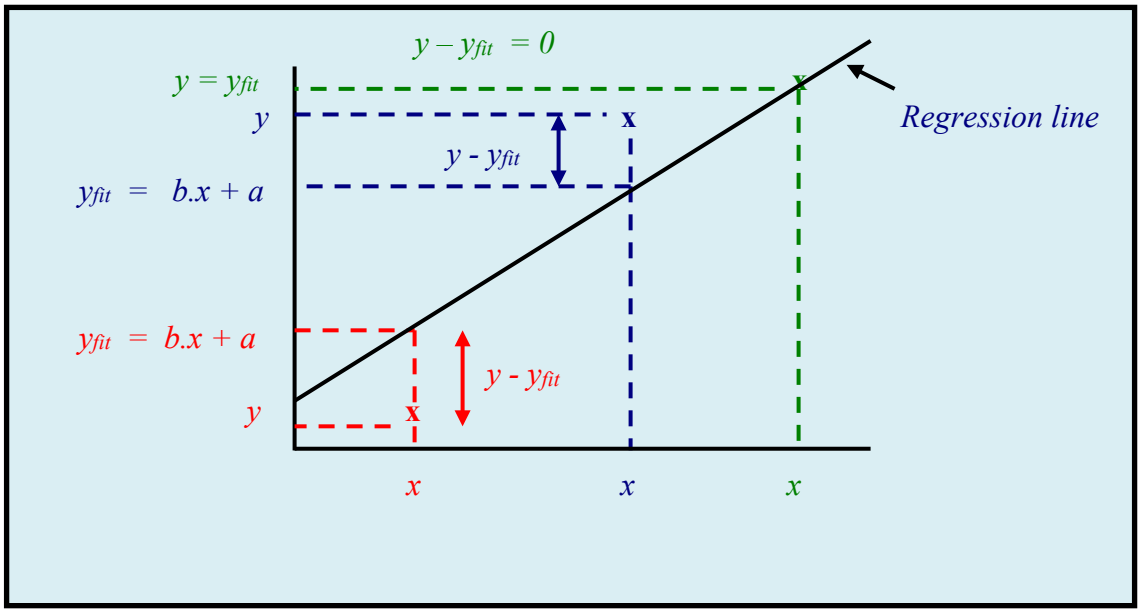


Figure 12.7 **Regression line determined for three values of the dependent variable (x,x and x) and the corresponding values for the independent variable (y,y and y). The regression line is drawn (from calculated values for slope b and intercept a) such that the sum of the squares of the residuals ($\Sigma(y - y_{fit})^2 = (y - y_{fit})^2 + (y - y_{fit})^2$) + $(y - y_{fit})^2$ is a *minimum***

As in other situations these positive and negative residuals cancel but the sum of their *squares* always gives a positive value which is a measure of the overall residuals of the observations from the line. The problem is to derive values for a and b such that $\sum(y - y_{fit})^2$ is a *minimum* (hence this calculation is also known as the *method of least squares*). Mathematicians deal with this problem by equating the two derivatives of $\sum(y - y_{fit})^2$ with respect to both a and b, to zero, then solving the resulting simultaneous equations for a and b. The solution for the slope of the line (b) is given by the expression:

$$b = \frac{\sum(x - m_x)(y - m_y)}{\sum(x - m_x)^2} \quad \text{..... Eq. 12.7}$$

The slope of the regression line (b) is also known as the *regression coefficient*. A similar expression can be derived to determine a. It is simpler, however, to use the fact that the line must pass through the intersection point of the means of x and y, so that equation 12.1 becomes $m_y = b.m_x + a$. Simple rearrangement, with substitution of the value for b enables determination of the value of a:

$$a = m_y - b.m_x \quad \text{..... Eq. 12.8}$$

The regression process assumes that the distribution of residuals ($y - y_{fit}$) about the regression line is Guassian i.e. that there are approximately equal numbers of observations each side of the line, that the residuals are independent of the value of x and that most observations are close to the line with relatively few a large distance form it. The standard deviation of the residuals (s_{res}, sometimes known as s_{yx}) is a valuable indicator (the lower the better) of the goodness of fit of the data to a straight line:

$$s_{res} = \sqrt{\frac{\Sigma (y - y_{fit})^2}{(n - 2)}} \quad \text{..... Eq.12.9}$$

which can be shown to be algebraically equal to:

$$s_{res} = \sqrt{\frac{(s_y)^2 (1 - r^2)(n - 1)}{(n - 2)}} \quad \text{..... Eq. 12.10}$$

Having obtained a value for the slope (b) the question often arises as to whether or not it is significant i.e. whether it is really different from zero (in which there would be no relationship between x and y). This question can be answered by dividing the difference between b and 0 (actually b) by the standard error of the estimate of b so as to obtain a corresponding value for t:

$$SE_b = \frac{s_{res}}{\sqrt{\Sigma(x - m_x)^2}} \quad \text{..... Eq 12.11}$$

$$t = \frac{b \sqrt{\Sigma(x - m_x)^2}}{s_{res}} \quad \text{..... Eq 12.12}$$

where t has n-2 degrees of freedom. If n is greater than 30 then the probability can be obtained from tables of z. Ninety-five percent confidence limits for the slope are given by $b \pm 1.96\ t$.

Question Q 12(1)

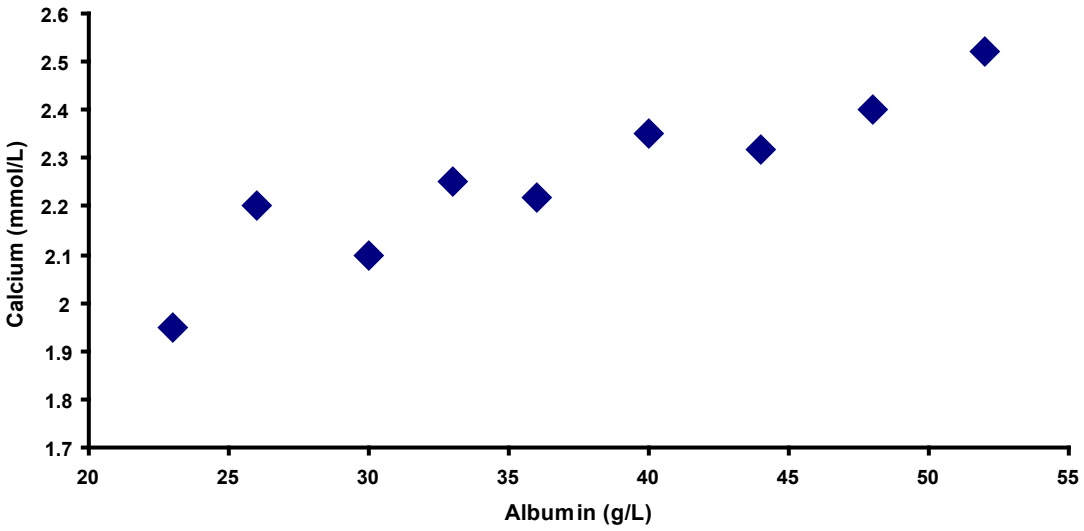
The following results for total calcium and albumin were obtained for a series of serum samples:

Sample	Albumin (g/L)	Calcium (mmol/L)
1	23	1.95
2	26	2.20
3	30	2.10
4	33	2.25
5	36	2.22
6	40	2.35
7	44	2.32
8	48	2.40
9	52	2.52

Is there a significant linear relationship between serum total calcium and albumin? Derive an expression to "correct" serum calcium to a "normal" albumin concentration of 40 g/L.

Answer Q 12(1)

First plot and inspect the data:



Inspection of the data suggests that it could be reasonably described by a linear expression (although it could be coming non-linear at the extremes). Construct a table with columns for x (albumin g/L), x^2, y (calcium mmol/L), y^2 and xy. Calculate the sum of each column i.e. Σx, Σx^2, Σy, Σy^2 and Σxy:

x	x^2	y	y^2	xy
23	529	1.95	3.803	44.85
26	676	2.20	4.840	57.20
30	900	2.10	4.410	63.00
33	1089	2.25	5.063	74.25
36	1296	2.22	4.928	79.92
40	1600	2.35	5.523	94.00
44	1936	2.32	5.382	102.08
48	2304	2.40	5.760	115.20
52	704	2.52	6.350	131.04
$\Sigma x = 332$	$\Sigma x^2 = 13034$	$\Sigma y = 20.31$	$\Sigma y^2 = 46.059$	$\Sigma xy = 761.54$

Calculate the correlation coefficient (r) using Eq 12.3:

$$r = \frac{\Sigma xy - (\Sigma x \Sigma y/n)}{\sqrt{\{[\Sigma x^2 - (\Sigma x)^2/n]\,[\Sigma y^2 - (\Sigma y)^2/n]\}}}$$

$$r = \frac{761.54 - (332 \times 20.31/9)}{\sqrt{\{[13034 - (332 \times 332/9)]\,[46.059 - (20.31 \times 20.31/9)]\}}}$$

$$= \frac{761.54 - 749.21}{\sqrt{\{[13034 - 12247]\,[46.059 - 45.833]\}}}$$

$$= \frac{12.33}{\sqrt{\{787 \times 0.226\}}} = \frac{12.33}{\sqrt{177.86}}$$

$$= \frac{12.33}{13.336} = \mathbf{0.925}$$

From tables of r, the probability of obtaining a value of 0.925 for n-2 degrees of freedom (7) is approximately 0.001 (i.e. 0.1%). Therefore the correlation is highly significant. (Alternatively t can be calculated using Eq 12.5 and its P value obtained from tables of t for 7 degrees of freedom.) Furthermore:

$$R^2 = 0.925^2 = \mathbf{0.856}$$

So that 85.6% of the variability of y (total calcium) can be explained by variation in x (albumin).

The slope of the regression line of y upon x can be obtained using Eq. 12.7:

$$b = \frac{\sum(x - m_x)\,(y - m_y)}{\sum(x - m_x)^2}$$

which can also be written:

$$b = \frac{\Sigma xy - (\Sigma x \Sigma y/n)}{\Sigma x^2 - (\Sigma x)^2/n}$$

Therefore $b = \frac{761.54 - (332 \times 20.31/9)}{13034 - (332 \times 332/9)} = \frac{12.33}{787} = \mathbf{0.0157\ mmol/g}$

The standard deviation of the residual (s_{res}) is calculated from Eq 12.10:

$$s_{res} = \sqrt{\frac{(s_y)^2 (1 - r^2) (n - 1)}{(n - 2)}}$$

$$(s_y)^2 = \frac{\Sigma(y - m_y)^2}{n - 1} = \frac{\Sigma y^2 - (\Sigma y)^2/n}{n - 1}$$

Therefore $$s_{res} = \sqrt{\frac{[\Sigma y^2 - (\Sigma y)^2/n] (1 - r^2) (n - 1)}{(n - 1)(n - 2)}}$$

Substituting for Σy^2, Σy, r and n:

$$s_{res} = \sqrt{\frac{[46.059 - (20.31^2/9)] (1 - 0.925^2)}{(9 - 2)}} = \mathbf{0.068}$$

The standard error of the slope (SE_b) is calculated from Eq 12.11:

$$SE_b = \frac{s_{res}}{\sqrt{\Sigma(x - m_x)^2}} = \frac{0.068}{\sqrt{[13034 - (332^2/9) =]}} = \mathbf{0.0024}$$

and the corresponding t value from Eq. 12.12

$$t = \frac{b \sqrt{\Sigma(x - m_x)^2}}{s_{res}} = \frac{0.0157}{0.0024} = \mathbf{6.5}$$

From tables when t = 6.5 for 7 degrees of freedom, P<0.001. Therefore the regression coefficient is significantly different from zero.

The value for the intercept (a) can be obtained by substituting for b, m_x and m_y in Eq. 12.8:

$$a = m_y - b.m_x = \Sigma y/n - b.\Sigma x/n$$

$$= 20.31/9 - 0.0157 \times 332/9 = 2.26 - 0.579 = \mathbf{1.68\ mmol/L}$$

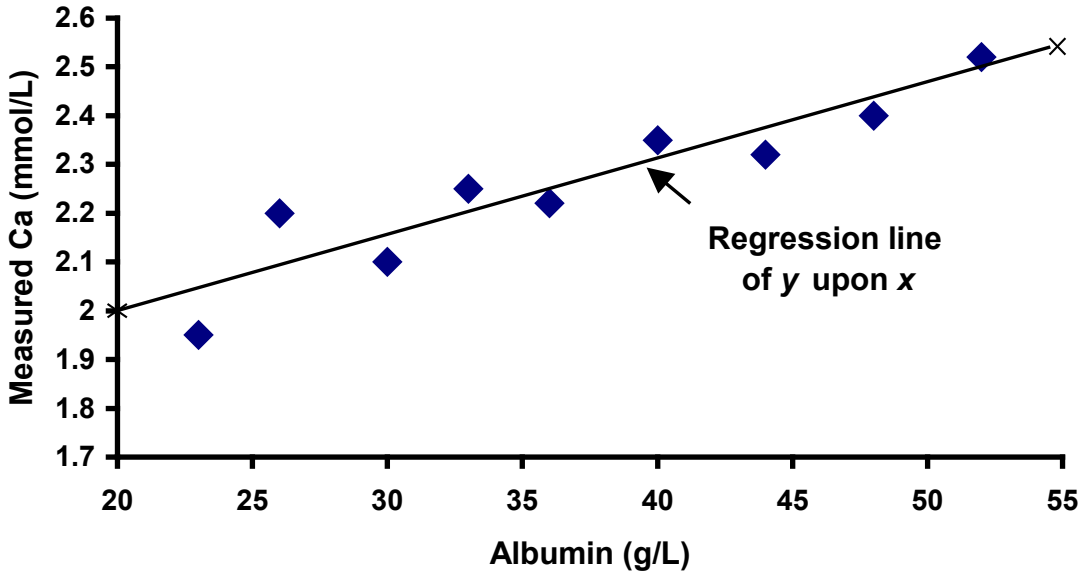
Therefore the equation for the regression line of y upon x is:

$$\boldsymbol{y = 0.0157\,x + 1.68}$$

To draw the regression line calculate the value of y for two carefully chosen values of x (say 20 and 55 g/L) using the regression equation, plot the points on a graph of y versus x and join them up:

when x = 20 g/L, y = (0.0157 x 20) + 1.68 = 1.99 mmol/L

when x = 55 g/L, y = (0.0157 x 55) + 1.68 = 2.54 mmol/L



The regression equation shows that for each increase in albumin by 1 g/L, the measured calcium also increases by 0.0157 mmol/L. To "correct" a measured calcium concentration to the value expected if the albumin was "normal" (40 g/L) the *difference* between the measured albumin and 40 g/L is multiplied by 0.0157 then *added* to the measured calcium:

Corrected Ca (mmol/L) = Measured calcium (mmol/L)

+ 0.0157 (40 – measured albumin, g/L)

If the measured albumin is greater than 40 g/L, then the expression 0.0157 (40 – measured albumin) becomes *negative* so that it is *subtracted* from the measured calcium.

Other modes of regression

Regression of x upon y. If it is y that is the independent variable then a regression of x upon y is performed. This involves minimizing the sum of the residuals between x and x_{fit} (Fig 12.3(b) rather than between y and y_{fit} (Fig 12.8(a)). The expression for the slope becomes:

$$b = \frac{\sum(x - m_x)(y - m_y)}{\sum(y - m_y)^2}$$

Sometimes it is not clear which is the independent variable. In this case a regression of both y upon x and of x upon y can be performed. Two different regression lines are obtained which intercept at the intersection of the two means (m_x and m_y).

Deming regression Ideally if neither x nor y can be identified as the independent variable then the best solution is to carry out a regression using the method of Deming. This involves calculating a regression line such that the sum of squares of residuals between the data points and lines drawn perpendicular to the regression line is minimal (see Fig 12.8 (c)). The process is rather complicated but computer programmes are available to perform the calculations.

Regression through the origin If the nature of the problem dictates that the regression line must pass through the origin then there is a simple technique to make sure this is so. Two points are needed to draw a straight line. The origin will be one and the other is at the intersection of the two means (see Fig 12.8 (d)).

$$b = \frac{m_y}{m_x} = \frac{\Sigma y/n}{\Sigma x/n} = \frac{\Sigma y}{\Sigma x}$$

Weighted regression The underlying assumption with linear regression is that the standard deviation of y is constant throughout the range of values. This is often not true and techniques exist for weighting each value to allow for variations in imprecision.

Multiple regression Sometimes we are dealing with more that one independent variable. In this situation the partial regression coefficients between pairs of variables are calculated. These techniques are beyond the scope of this book.

Non-linear regression Although non-linear data can often be transformed then analysed by linear regression this is not always the case. Non-linear methods are available but again are beyond the scope of this book.

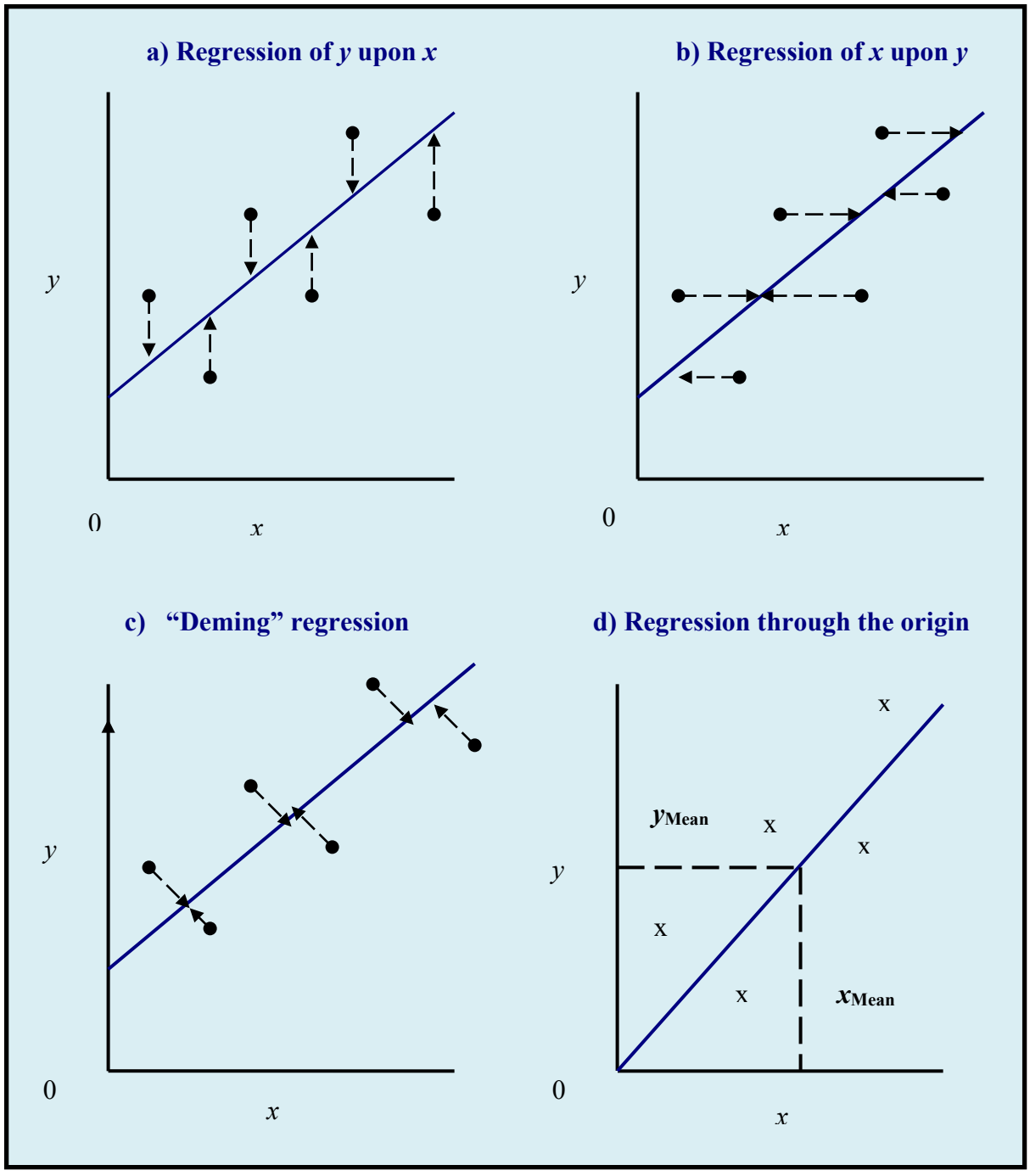


Figure 12.8 **Modes of regression. In each case the sum of the squares of the distance from the line (⟶) is minimized in the direction shown**

Method comparison – a special case?

Method comparisons are usually carried out to determine if two methods yield the same results with patient samples (o rare sufficiently similar for routine purposes). The correlation coefficient is unhelpful since it only indicates if there is a relationship between the two sets of results, not whether the results are comparable; furthermore it would be surprising if the correlation was not significant between two methods designed to measure the same analyte. Simple regression methods rely on the assumption that the independent variable is determined without error – an assumption which is rarely true. Regression methods do however, provide the relationship between the two methods.

Since two methods are measuring the same analyte, the goal is determine whether or not the two sets of results are significantly different. One way to do this is by the paired *t*-test (see Chapter 11). However, if the two sets of results are identical then the relationship between them should be described by a straight line which passes through the origin and has a slope of one (i.e. the data is best described by the expression $y = x$). Visual assessment of the data can be made if the two sets of results (x and y) are plotted against each other and the line $y = x$ drawn (Fig 12.9). *Altman and Bland* took this approach one step further and plotted the absolute difference d (equal to $y - x$) against the mean of each pair of data, $(x + y)/2$. If there is no significant difference between the two methods then the data should scatter evenly about a horizontal line with a value of zero (Fig 12.9); in other words the mean difference (m_d) should be close to zero.

$$m_d = \frac{\Sigma d}{n} \quad \text{..... Eq. 12.13}$$

Calculation of the standard deviation of the differences gives a measure of the scatter about the line.

$$s_d = \sqrt{[\Sigma (d - m_d)^2/(n - 1)]} \quad \text{..... Eq. 21.14}$$

The 95% confidence limits of the mean difference are known as the "95% *limits of agreement":*

$$\text{95\% } \textit{limits of agreement} = m_d - 1.96s_d \text{ to } m_d + 1.96s_d \quad \text{..... Eq. 21.15}$$

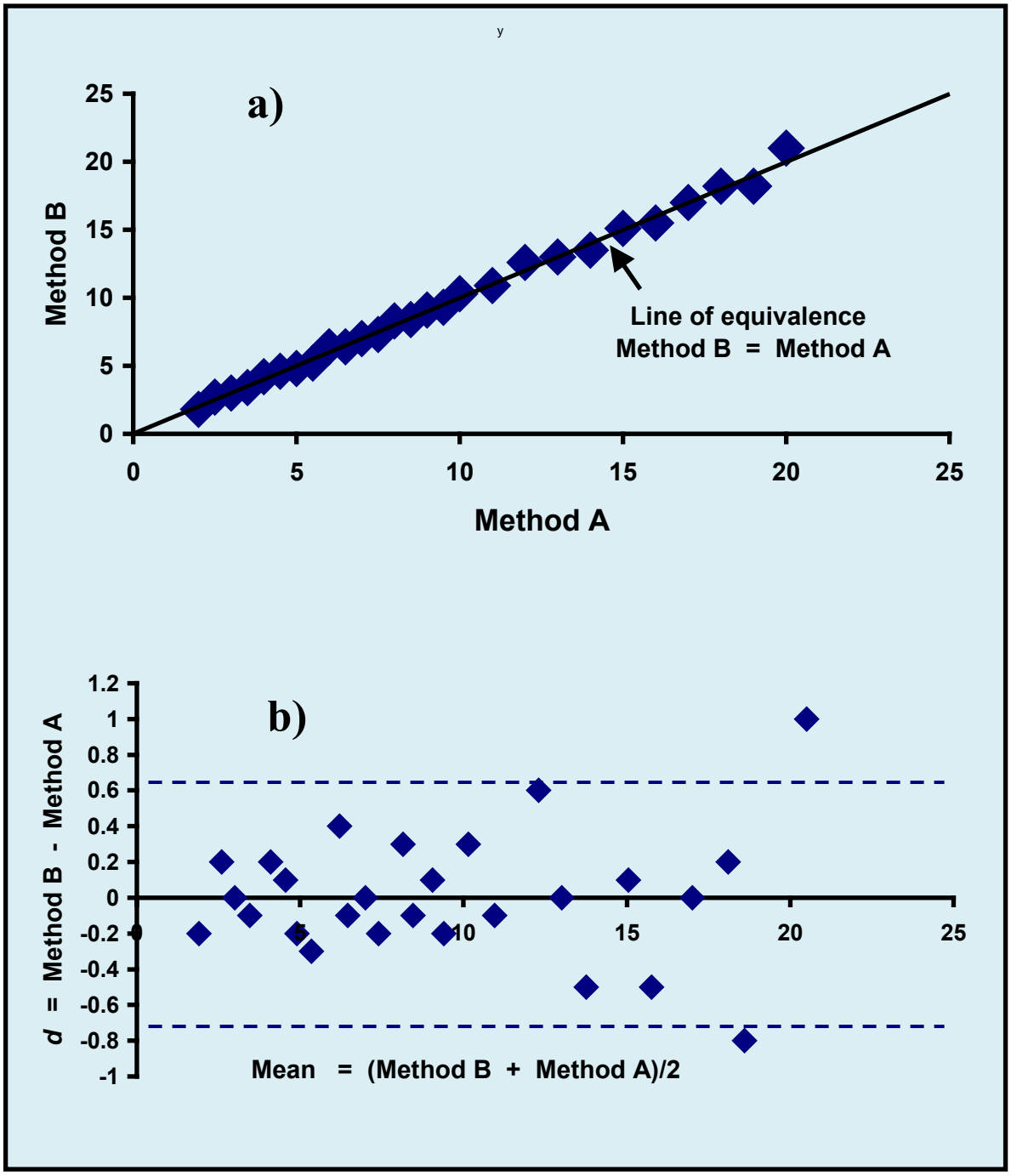


Figure 12.9 **Data from a comparison of two methods – A and B. a) plot of Method B versus Method A results, b) difference plot (Altman & Bland) of same data with 95 % limits of agreement shown**

A *t*-test can be applied to determine if the mean difference is significantly different from zero:

$$t = \frac{m_d}{s_d/\sqrt{n}} \quad \text{..... Eq. 12.16}$$

These calculations are identical to those of the paired *t*-test (Chapter 11).

It is important to carefully inspect the data to ensure that the differences are normally distributed about their mean throughout the range of results. For example, if there is a significant slope in the *y* versus *x* plot of the data then the difference plot will show an even more exaggerated change in *d* with concentration, which can be removed if the percentage difference is plotted instead of the absolute difference.

The advantage of the Altman-Bland approach is that it focuses on the differences of individual results throughout the range. The same information can be obtained from plots of *y* versus *x*, but small deviations are not obvious.

ADDITIONAL QUESTIONS

1. Regression analysis of results using new standards (*y*) against old standards (*x*) showed a linear relationship. The regression coefficient (slope) was 1.10 and the intercept on the *y* axis 1.0 mmol/L. Calculate the results which would be expected using new standards for the analysis of old standards containing (a) 15 mmol/L and (b) 150 mmol/L.

2. A laboratory changed its method for the assay of serum alkaline phosphatase activity. Assay of a selection of patient's samples by both methods yielded the following data:

ALP (Old method), IU/L:	50	350	700	100	1500	2000	420	1200
ALP (New method), IU/L:	40	190	350	90	750	1500	280	600

   A gastroenterologist has been using ALP to monitor patients on treatment. Use these data to derive an expression to convert the new ALP results to the results expected by the old method.

3. An endocrinologist has been using serum prolactin measurements to assess the response of patients with prolactinoma to treatment with a new drug. The following data were obtained for a series of patients:

Drug dosage (mg/kg body wt):	50	100	150	200	250	300	350	400
Prolactin (IU/L)	750	1500	350	400	2000	1250	500	1800

Do these data show a linear relationship between drug dosage and serum prolactin concentration?

4. A research paper contains the following statement:

"A good correlation was obtained when 45 patient samples were analysed by methods A and B ($r = 0.90$, $B = 1.05A - 10$)…." Comment on this statement.

Chapter 13

Clinical utility of laboratory tests

A laboratory test for a specific disease may yield either a positive or a negative result. This may be achieved either by a qualitative test (in which the result is either positive or negative) or a quantitative result which is either above (positive) or below (negative) a selected "cut-off" value. In the ideal world all patients with the disease in question, and only these patients, would give a positive result. Furthermore, all patients without the disease, and only these patients, would produce negative results. In other words there would be no false positives or false negatives. However, in reality a proportion of both false positives (positive results in patients without disease) and false negatives (negative results in patients with disease) are always obtained. The classification of positive and negative results is shown in Fig 13.1. Correct interpretation of laboratory results demands an appreciation of the likelihood of the result identifying the presence (or absence) of disease.

	Positive result	**Negative result**	**Total**
Patients *with* disease	**TP**	**FN**	**TP + FN**
Patients *without* disease	**FP**	**TN**	**TN + FP**
Total	**TP + FP**	**TN + FN**	**TP + FN + TN +FP**

Figure 13.1 Classification of positive and negative test results.
TP = true positive; FP = false positive
TN = true negative; FN = false negative

Sensitivity and specificity

The *sensitivity* of a test is the *proportion* of patients *with* disease that are identified by the test. The number with disease giving a positive result is the true positives (TP) and the total number with disease is the sum of the true positives and false negatives (TP + FN):

$$\textbf{Sensitivity} = \frac{\textbf{TP}}{\textbf{TP + FN}} \quad \textbf{..... Eq. 13.1}$$

The *specificity* of a test is the *proportion* of patents *without* disease that are identified by the test. The number without disease giving a negative result is the true negatives (TN) and the total number without disease is the sum of the true negatives and false positives (TN + FP):

$$\textbf{Specificity} = \frac{\textbf{TN}}{\textbf{TN + FP}} \quad \textbf{..... Eq. 13.2}$$

Sometimes these values are multiplied by 100 to give results for sensitivity and specificity as percentages.

The efficiency of a test is the proportion of all results which are true results:

$$\textbf{Efficiency} = \frac{\textbf{TP + FN}}{\textbf{total tested}} = \frac{\textbf{sensitivity + specificity}}{\textbf{2}} \quad \textbf{..... Eq. 13.3}$$

Determinations of sensitivity and specificity, like any other biological measurement, are subject to error. Consider a test with a true sensitivity of 0.50 (50%) determined on a group of 50 patients with a particular disease. If the sensitivity was repeatedly determined on groups of 50 patients with the disease then the sensitivity would not come out at exactly 0.5 every time but would cluster around this value in a similar way to the mean determined for a continuous variable. Situations where there are only two possible outcomes (positive or negative) belong to a special distribution called the *binomial distribution*. Fortunately if the sample size (n) is large enough (>30) then the binomial approximates to a normal distribution, and this can be extremely useful.

If p is the *proportion* of n individuals with the disease giving a positive *result* (i.e. estimated sensitivity), then the variance, standard deviation and 95% confidence limits of p are calculated as follows.

$$s^2 = p\,(1-p)\,n \quad \text{..... Eq. 13.4}$$

$$s = \sqrt{[\,p\,(1-p)\,n]} \quad \text{..... Eq. 13.5}$$

$$\text{95\% confidence limits} = p + 1.96\,s \text{ to } p - 1.96\,s \quad \text{..... Eq. 13.6}$$

Question Q 13(1)

A test for a certain disease gave a 5% false positive rate and a 2% false negative rate. What is the sensitivity and specificity of the test?

Answer Q13(1)

If the false negative rate is 2% then this means that for every 100 patients with disease, 2 false negatives will be obtained, the remainder will give positive results.

Therefore FN = 2 and TP = 100 – 2 = 98 and (TP + FN) = 2 + 98 = 100

$$\text{Sensitivity} = \frac{TP}{TP + FN} = \frac{98}{100} = \mathbf{0.98} \text{ (or 98\%)}$$

If the false positive rate is 5% then this means that for every 100 patients without disease 5 false positive results will be obtained, the remainder will give negative results.

Therefore FP = 5 and TN = 100 – 5 = 95 and (TN + FP) = 5 + 95 = 100

$$\text{Specificity} = \frac{TN}{TN + FP} = \frac{95}{100} = \mathbf{0.95} \text{ (or 95\%)}$$

Predictive values

Sensitivity and specificity define the performance of a test when applied to populations of individuals who either have (sensitivity) or do not have (specificity) the disease in question. In practice tests are applied to populations made up of a mixture of subjects with or without disease. The proportion of the two populations (i.e. the p*revalence* of disease in the population being tested) can have a profound effect on the *predictive value* of a test. The predictive value of a test is the probability that a subject with a positive result has the disease (*positive* predictive value, denoted PV+) or the probability that a subject with a negative result does not have the disease (*negative* predictive value denoted PV-).

Positive predictive value (PV+) $= \frac{TP}{(TP + FP)}$ **…… Eq. 13.7**

Negative predictive value (PV-) $= \frac{TN}{(TN + FN)}$ **…….. Eq. 13.8**

Prevalence of disease $= \frac{(TP + FN)}{(TP + FN + TN + FP)}$ **……. Eq. 13.9**

Consider the example in question Q 13.1 where there is a false positive rate of 5%. In a population in which a half of all individuals have disease (i.e. the prevalence of disease is 0.5 or 50%) it is possible to calculate the number of each possible outcome of the test given that the sensitivity is 0.98 and specificity 0.95:

If prevalence is 0.5, then proportion of patients with disease (TP + FN) = 0.5.

Since sensitivity $= \frac{TP}{TP + FN}$, TP = sensitivity (TP + FN) = 0.98 x 0.5 = 0.49

In other words:

TP = Sensitivity x Prevalence …….. Eq. 13.10

If sensitivity is expressed as a percentage then the prevalence is multiplied by sensitivity and divided by hundred i.e. in the above example the sensitivity is 85% so that the prevalence is multiplied by 85/100.

Similarly if prevalence is 0.5, then proportion without disease (TN + FP) = 0.5

If the prevalence is expressed as a proportion of one rather than absolute numbers or a percentage then (TN + FP), also expressed as proportion of one, is obtained by subtracting the prevalence from one:

(TN + FP) = 1 - prevalence ..... Eq. 13.11

Since specificity = $\frac{TN}{TN + FP}$, TN = specificity (TN + FP) = 0.95 x 0.5 = 0.475

In general:

TN = Specificity x (1 – prevalence) ..... Eq. 13.12

Since (1 - prevalence) = (TN + FP) = 0.5 and TN = 0.45, FP can be calculated:

FP = (1 - prevalence) - TN = 0.5 - 0.475 = 0.025

Since TP = 0.49 and FP = 0.025 the proportion of patients with a positive result who have disease (i.e. the positive predictive value) is

$$\frac{0.49}{(0.49 + 0.025)} = 0.95 \text{ (or 95\%)}$$

Therefore 0.95 (95%) of positive results are due to disease.

Consider a population consisting of 10 patients with and 90 patients without, disease (i.e. prevalence = 0.1 = TP + FN).

As above, TP = sensitivity (TP + FN) = 0.98 x 0.1 = 0.098

Since we are dealing with proportions all groups must add up to one

i.e. TP + FN + TN + FP = 1, and (TP + FN) = 0.1

then TN + FP = 1 - (TP + FN) = 1 - 0.1 = 0.9

Therefore TN = specificity (TN + FP) = 0.95 x 0.9 = 0.855

and FP = (TN + FP) - TN = 0.90. - 0.855 = 0.045

The proportion of true positives has fallen from 0.49 to 0.098 and false positives risen from 0.025 to 0.045. Therefore, the proportion of positive results which are due to disease (i.e. the positive predictive value) is:

$$PV+ = \frac{0.098}{(0.098 + 0.025)} = \frac{0.098}{0.143} = 0.685$$

Therefore only 0.685 (or 68.5%) of positive results are due to the presence of disease. Continuation of this process for other prevalences yields the data in Fig 13.2.

Prevalence	TP	FP	PV+
1 in 2	**4,900**	**250**	**0.95**
1 in 4	**2,450**	**375**	**0.87**
1 in 10	**980**	**450**	**0.69**
1 in 40	**245**	**488**	**0.33**
1 in 100	**98**	**495**	**0.17**
1 in 1,000	**10**	**499**	**0.02**
1 in 10,000	**1**	**500**	**0.002**

Figure 13.2 Effect of disease prevalence upon the number of true positives (TP), false positives (FP) and positive predictive value (PV+) for a test with a sensitivity of 98% and specificity of 95% applied to a population of 10,000 subjects

Therefore when the prevalence is very low the number of false positives exceed the true positives.

Question Q13(2)

In a cancer clinic where the prevalence of ovarian malignancy is 40%, a tumour marker has a specificity of 88% and a sensitivity of 92%. Calculate the predictive value of a positive test result.

Question Q13(2)

To solve this type. of problem the calculations can be based on actual numbers of results, percentages or proportions of one. It is often simplest to work with proportions.

Construct a 2 x 2 contingency table (as Fig 13.1):

	Positive result	**Negative result**	**Total**
Patients *with* disease	TP	FN	Prevalence
Patients *without* disease	FP	TN	1 - prevalence

The prevalence of disease is 40 % (0.4 as a proportion) so that (1 – prevalence) is (1 – 0.4) = 0.6. The above table then becomes:

	Positive result	**Negative result**	**Total**
Patients *with* disease	TP	FN	0.4
Patients *without* disease	FP	TN	0.6

The next task is to determine values for TP, FN, FP and TN. We are given values for sensitivity and specificity so can write:

$$\text{Sensitivity} = \frac{TP}{TP + FN} = 0.92 \quad \text{and} \quad \text{specificity} = \frac{TN}{TN + FP} = 0.88$$

Since the prevalence is 0.4 and is equal to (TP + FN), and (1 – prevalence) is 0.6 and is equal to (TN + FP), both of these expressions can be re-written:

$$\text{Sensitivity} = \frac{TP}{0.4} = 0.92 \quad \text{and} \quad \text{specificity} = \frac{TN}{0.6} = 0.88$$

These expressions are then rearranged and solved for TP and TN:

$$TP = 0.92 \times 0.4 = 0.368 \quad \text{and} \quad TN = 0.88 \times 0.6 = 0.528$$

Both of these values are now added to the contingency table:

	Positive result	**Negative result**	**Total**
Patients *with* disease	0.368	FN	0.4
Patients *without* disease	FP	0.528	0.6

Since prevalence = (TP + FN) and (1 – prevalence) = (TN + FP), values for FN and FP can be obtained by subtraction of TP and TN from the corresponding totals in the last column.

	Positive result	**Negative result**	**Total**
Patients *with* disease	0.368	0.032	0.4
Patients *without* disease	0.072	0.528	0.6

The predictive value of a positive test (PV+) is then obtained by substitution of TP and FP into equation Eq 13.7:

$$PV+ = \frac{TP}{TP + FP} = \frac{0.368}{0.368 + 0.072} = \frac{0.368}{0.440} = \mathbf{0.84}\ (84\%)$$

Question Q13(3)

A man has a PSA of 5 µg/L. 22% of patients with benign prostatic hypertrophy and 38% of patients with prostatic cancer have concentrations of PSA between 4.1 and 10 µg/L. What is the positive predictive value for a diagnosis of cancer of the result for this man in this range, if the prevalence of cancer in his age group is 5% and benign prostatic hypertrophy is 20%? Assume 2% of patients without any prostatic pathology have a PSA >4.1 µg/L.

Answer Q13(3)

This question differs from *Q13(2)* in that there are three instead of two groups of patients. However, there are only two groups as far as the disease in question (prostatic cancer) is concerned – those with cancer and those without. The only difference is that the group without cancer is made up of two populations:

- those with benign prostatic hypertrophy (BPH)
- those *without* BPH *and* without prostatic cancer (CAP)

In order to calculate the predictive values first calculate the individual positive and negative results for each of the three groups. This it is easier if everything is converted to proportions of one rather than working with percentages.

CAP group, prevalence = 5% = 5/100 = 0.05

BPH group, prevalence = 20% = 20/100 = 0.20

The normal group (i.e. those with neither BPH or CAP) consists of the remaining patients

Therefore, prevalence of normals = 1.0 – (0.05 + 0.20) = 0.75

Sensitivity_{CAP} = 38% = 38/100 = 0.38

Sensitivity_{BPH} = 22% = 22/100 = 0.22

Set up a contingency table and enter the totals:

	Positive result	**Negative result**	**Total**
CAP	TP_{CAP}	FN_{CAP}	0.05
BPH	TP_{BPH}	FN_{BPH}	0.20
Normals	FP	TN	0.75

For the group with CAP:

$$TP_{CAP} = Prevalence_{CAP} \times Sensitivity_{CAP}$$

$$= 0.05 \times 0.38 = 0.019$$

Since TP_{CAP} and FN_{CAP} must add up to $prevalence_{CAP}$

$$FN_{CAP} = 0.05 - 0.019 = 0.031$$

Similarly for the group with BPH:

$$TP_{BPH} = 0.20 \times 0.22 = 0.044$$

and $$FN_{CAP} = 0.20 - 0.044 = 0.156$$

For the group without either disease we are told that 2% of patients have raised PSA (i.e. false positives). Therefore FP = 2% = 2/100 = 0.02

Therefore, TN = 0.75 - 0.02 = 0.73

Enter these values to complete the contingency table:

	Positive result	**Negative result**	**Total**
CAP	0.019	0.031	0.05
BPH	0.044	0.156	0.20
Normals	0.020	0.730	0.75
Total	0.083	0.917	1.00

The positive predictive value (PV+) is obtained by substituting the proportion of true positives and false positives into Eq 13.7, remembering that the false positives are made up of two groups – those with BPH and those with neither BPH nor CAP:

$$PV+ = \frac{TP_{CAP}}{(TP_{CAP} + TP_{BPH} + FP)}$$

$$= \frac{0.019}{(0.019 + 0.044 + 0.02)} = \mathbf{0.23} \text{ (2 sig figs) or 23\%}$$

Therefore approximately only 1 in 4 positive results will be due to prostatic cancer.

The positive predictive value (the likelihood of disease when a test result is positive) is markedly influenced by the prevalence of disease in the population being tested. The prevalence of any disease is very low in the general population, becomes higher in a population with symptoms of the disease in question and highest in patients referred by a general practitioner to a hospital specialist.

This concept is of greatest significance when screening a well population for the presence of a rare disease such as bowel cancer. False positives may be obtained due to interference in the screening test used, dietary factors and bleeding from other sources. As a result the predictive value of a positive screening test is low so that only a small portion of patients with positive results will have bowel cancer. Detection of bowel cancer then relies on a secondary test (such as colonoscopy). However, the screening test is still valuable in that it enables selection of a sub-population of individuals with a higher prevalence of disease which warrants the expensive and unpleasant secondary test. On the other hand a false negative will be obtained if the tumour is not bleeding when the screening test is carried out. When designing a screening program the incidence of false positives and false negatives must be carefully balanced so as to achieve optimal detection of disease with minimum cost and morbidity.

Receiver operator characteristic (ROC) curves

Although calculations of sensitivity, specificity etc are simple with a qualitative test which yields either a positive or negative result, the majority of diagnostic tests are of a quantitative nature yielding a result which is a continuous variable. The quantitative result is converted into a qualitative one by using a decision level (or cut-off point). Values above this level are classified as "positive" and below it as "negative". This provides an opportunity to manipulate the performance of a test by adjustment of the decision level.

Figure 13.3a shows the result for an analyte which is able to distinguish patients with a particular disease from healthy individuals. The distribution of results for each set of patients is roughly Guassian, but the two distributions overlap. In this overlap area it is impossible to use the test to distinguish patients with the disease from those who have not. If a decision level at point "A" is used then there are no false negatives but as many as a half of those without disease would yield false positive results (i.e. the sensitivity would be virtually one but with a specificity of only 0.5). A decision level of "B" produces fewer false positives but at the expense of missing a small number with disease. Using "C" or "D" even fewer false positives are obtained but false negatives become more frequent. At level "E" there are no false positives but at the expense of missing about one half of patients with the disease.

A useful way of looking at the effect of changing decision levels is to plot the sensitivity (*y*-axis) versus 1 – specificity (*x*-axis) at each decision level (see Fig 13.3b). The *x*-axis therefore represents the false positive rate and the *y*-axis the true positive rate. For a perfect test, the resulting *receiver operator characteristic* (*ROC*) curve would extend from the lower left to the upper left then to the upper right. It is generally accepted that for a test of no diagnostic value the curve would be a diagonal line from the origin with a slope of one. When comparing the performance of different tests it is helpful to calculate the area under the *ROC* curve. Other things being equal, the test with the highest area is superior.

Although examination of the ROC curve is often helpful, usually other considerations (the importance of avoiding false positives versus false negatives, the nature and cost of any follow-up tests etc) determine which decision point to use.

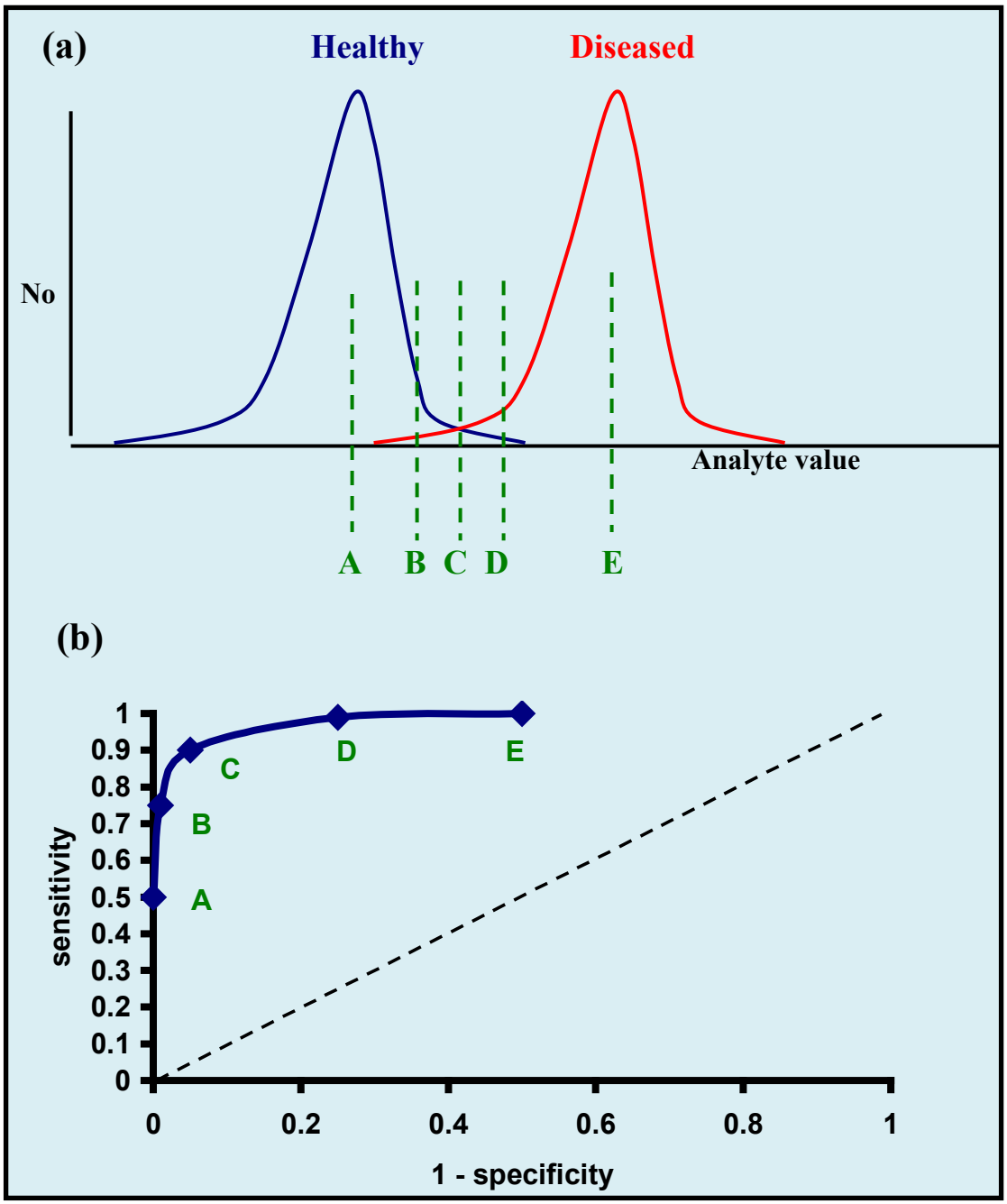


Figure 13.3 **Effect of variation of decision level (A,B, C, D and E) upon the performance of a diagnostic test. a) Distribution of results from healthy and diseased individuals; b) Same data plotted as a receiver operator characteristic (ROC) curve**

Odds and likelihood ratios

What the clinician really needs to know is the probability that a patient with a given test result has the disease in question. To answer this question two parameters are required:

- The prevalence of the disease in the population to which the patient belongs
- The likelihood ratio of a positive test

The *likelihood ratio positive* (LR+) is defined as the ratio between the probability of finding a positive test in the presence of disease and the probability of a positive test in the absence of disease. The probability of a positive test in the presence of disease is simply the sensitivity of the test. The probability of finding a negative test in the absence of disease is the specificity so that the probability of a positive test when disease is absent is simply (1 – specificity). Therefore LR+ can be calculated directly from the sensitivity and specificity of the test:

$$\textbf{LR+} = \frac{\textbf{probability of +ve test with disease}}{\textbf{probability of +ve test without disease}} = \frac{\textbf{sensitivity}}{\textbf{(1 – specificity)}} \quad \textbf{… Eq. 13.13}$$

A similar expression can be derived for likelihood ratio of a negative test (LR-):

$$\textbf{LR-} = \frac{\textbf{probability of a –ve test with disease}}{\textbf{probability of a -ve test without disease}} = \frac{\textbf{(1 – sensitivity)}}{\textbf{specificity}} \quad \textbf{… Eq. 13.4}$$

When comparing several different tests it is often helful to calculate the *diagnostic odds ratio*, which is simply the ratio of LR+ to LR-, which can be simplified to:

$$\textbf{Diagnostic odds ratio} = \frac{\textbf{sensitivity x specificity}}{\textbf{(1 – sensitivity) x (1 – specificity)}} \quad \textbf{……. Eq. 13.15}$$

The likelihood ratio positive is useful in converting the *pre-test odds* (the likelihood that disease is present before the test is carried out) to the *post-test odds*. In the absence of any data about the patient's individual risk factors, the pre-test odds is the prevalence of the particular disease in the population to which he/she belongs. For example if the prevalence of disease is 0.2 (20%) then a group of 100 patients will contain 20 with disease and 80 without disease. Therefore the odds of disease being present are given by the ratio of the number of patients with disease to the number without i.e. 20/80 = 0.25. In other words for each patient without disease there will 0.25 patients with disease or rather for every 4 patients without disease there will be one with the disease or the odds *against* disease being present are 4:1. In general:

$$\textbf{Pre-test odds} = \frac{\textbf{prevalence}}{\textbf{(1 – prevalence)}} \quad \textbf{..... Eq. 13.16}$$

provided the prevalence is expressed as an absolute proportion of one

Multiplication of the pre-test odds by the likelihood ratio positive gives the post-test odds:

Post-test odds = pre-test odds x likelihood ratio positive..... Eq. 13.17

Post-test odds can be converted back to the probability of the patient having disease using the expression:

$$\textbf{Post-test probability} = \frac{\textbf{post-test odds}}{\textbf{(1 + post-test odds)}} \quad \textbf{..... Eq. 13.18}$$

Provided reliable data is available on the prevalence of the disease and the diagnostic performance of the laboratory test then calculation of post-test probability is the best way of presenting laboratory data to the clinician. However, one problem that is not addressed is the effect of the magnitude of the increase in the analyte above the decision point. A result which is several times the decision value often is more likely to be associated with disease than a result which is only just above this value.

Two further refinements of this technique are often used:

- Multiple tests. Provided the likelihood ratios of several different tests are known then it sometimes possible to combine them to give a more reliable post-test probability of disease than for either test on its own. An example of this is the triple test to screen for Down's syndrome.

- Sequential testing. A preliminary laboratory test may be applied to select a population in which the prevalence of disease is enhanced, then apply a secondary test to accurately identify those with the disease. In is vital that the two tests are independent i.e. are not different ways of assessing the same thing.

Question Q 13(4)

A certain disease has a prevalence of 10 percent. A diagnostic test was applied to a random sample of 500 individuals from this population and yielded 45 true positive and 40 false positive results. Calculate a) the likelihood ratio positive and, b) the post-test probability of disease being present for a positive test result.

Answer Q 13(4)

If the prevalence is 10% then for a sample of 500 individuals 50 will have the disease and 450 will be healthy. Use this data to set up a 2x2 contingency table:

	Positive test	Negative test	Total
Diseased	45	FN	50
Healthy	40	TN	450

Subtract the positive test from the corresponding total to give to give the numbers of negative results:

	Positive test	Negative test	Total
Diseased	45	5	50
Healthy	40	410	450

Use these values to calculate the sensitivity and specificity:

$$\text{Sensitivity} = \frac{TP}{(TP + FN)} = \frac{45}{50} = 0.90$$

$$\text{Specificity} = \frac{TN}{(TN + FP)} = \frac{410}{450} = 0.911$$

a) Calculate the likelihood ratio positive (LR+) from the sensitivity and specificity (Eq13.13):

$$LR+ = \frac{\text{sensitivity}}{(1 - \text{specificity})} = \frac{0.90}{(1 - 0.911)} = \frac{0.90}{0.089} = \mathbf{10} \text{ (2 sig figs)}$$

b) First calculate the pre-test odds from the disease prevalence (Eq 13.16):

$$\text{Pre-test odds} = \frac{\text{prevalence}}{(1 - \text{prevalence})}$$

The prevalence is converted to a proportion of one by dividing the number with disease by the total number tested.

$$\text{Prevalence} = \frac{50}{500} = 0.1$$

$$\text{Pre-test odds} = \frac{0.1}{(1 - 0.1)} = \frac{0.1}{0.9} = 0.111 \quad \text{(i.e. 0.111 to 1)}$$

The post-test odds is calculated by multiplication of the pre-test odds by the likelihood ratio positive (Eq. 13.17):

$$\text{Post-test odds} = \text{Pre-test odds} \times LR+$$

$$= 0.111 \times 10 = 1.11$$

The post-test odds can be converted to a probability using Eq 13.18:

$$\text{Post test probability} = \frac{\text{post test odds}}{(1 + \text{post test odds})}$$

$$= \frac{1.11}{(1 + 1.11)} = \frac{1.11}{2.11} = \mathbf{0.53} \text{ (2 sig figs)}$$

ADDITIONAL QUESTIONS

1. A test for a particular disease has a sensitivity of 95% and a specificity of 95%. Calculate the predictive value of both a positive and a negative test result in a population in which the prevalence of the disease is:

   a) 1 in 2
   b) 1 in 5000

2. The table shows data from two urinary screening tests for the detection of phaeochromocytoma.

Test	Sensitivity	Specificity
VMA	96.7%	99.1%
Total metanephrines	100%	98%

   Both tests were used to screen a population of 100,000 hypertensive patients in which the incidence of phaeochromocytoma is known to be 0.5%.

   a) How many patients with phaeochromocytoma were missed by the VMA test?

   b) How many patients were incorrectly diagnosed as having phaeochromocytoma using the metanephrine test?

   c) Which test would you use to screen a hypertensive population for phaeochromocytoma? Give reasons for your choice.

3. A new laboratory test has a sensitivity of 85% and a specificity of 90%. The incidence of disease in a population considered at risk is 0.10. What is the predictive value of

   a) a positive result?
   b) a negative result?

4. A proposed diagnostic serological test for coeliac disease was evaluated in 200 consecutive patients referred to a paediatric gastroenterology service in whom the condition was suspected clinically. The test result was compared with the diagnosis as established by biopsy, withdrawal of gluten and response to re-challenge. On this basis 76 children had the condition of whom only 64 gave a positive test result: 10 positive test results occurred in children who were shown not to have coeliac disease. Calculate the sensitivity and specificity of the test and the predictive value of a positive result.

5. In a cancer clinic where the prevalence of ovarian malignancy is 40%, a tumour marker has a specificity of 88% and a sensitivity of 92%. Calculate the predictive value of a positive test result. If this test was used as a screening tool in all patients attending a general gynaecological clinic with a cancer prevalence of 0.4%, what would be the predictive value of a positive test in this population?

6. A certain disease has a prevalence of 5 percent. A diagnostic test was applied to a random sample of 400 individuals from this population and yielded 15 true positive and 30 false positive results. Calculate: a) the positive predictive value of the test applied to this population, b) the pre-test odds of disease, c) the likelihood ratio positive; d) the post test odds of disease for a positive result, and e) the post-test probability of disease for a positive result.

7. A two-stage sequential test strategy is used to screen for a rare inherited disease. The prevalence of the disease is 0.0005. The initial test has a sensitivity of 98% and specificity of 95%, the follow-up test a sensitivity of 95% and specificity of 99%. What is the probability of a patient with a positive result for the follow-up test having the disease?

Chapter 14

Statistical power

Power is the ability of a statistical test to detect a specified effect with a given probability. For example, comparison of the means of two sets of data involves calculation of the number of standard errors by which they differ (i.e. the z or t value) then looking up the P value of this statistic in appropriate tables. Fig 14.1 shows the sampling distribution of the mean for two sets of data. In Fig 14.1a it is apparent that the two distributions overlap considerably. However, increasing the sample size from 10 to 50 lowers the standard deviation of these distributions (i.e. the standard error of the means) so that the distributions are less widely spread but without any shift of the overall mean values. This is exactly what would be predicted from Fig 11.1. As a consequence the overlap between the distributions has been eliminated. Since the standard error of each mean is given by $s/\sqrt{n}$ it is not surprising that increasing the value of n reduces the standard error (i.e. spread) of each distribution so reducing the degree of overlap.

The likelihood of demonstrating a significant difference between the means of two sets of data increases as the degree of overlap is reduced. The degree of overlap and hence the chance of detecting a difference (i.e. power) is in turn determined by four factors:

- The sample size of each group.
- The magnitude of the difference between the two groups
- The standard deviation of the data
- The required level of statistical significance.

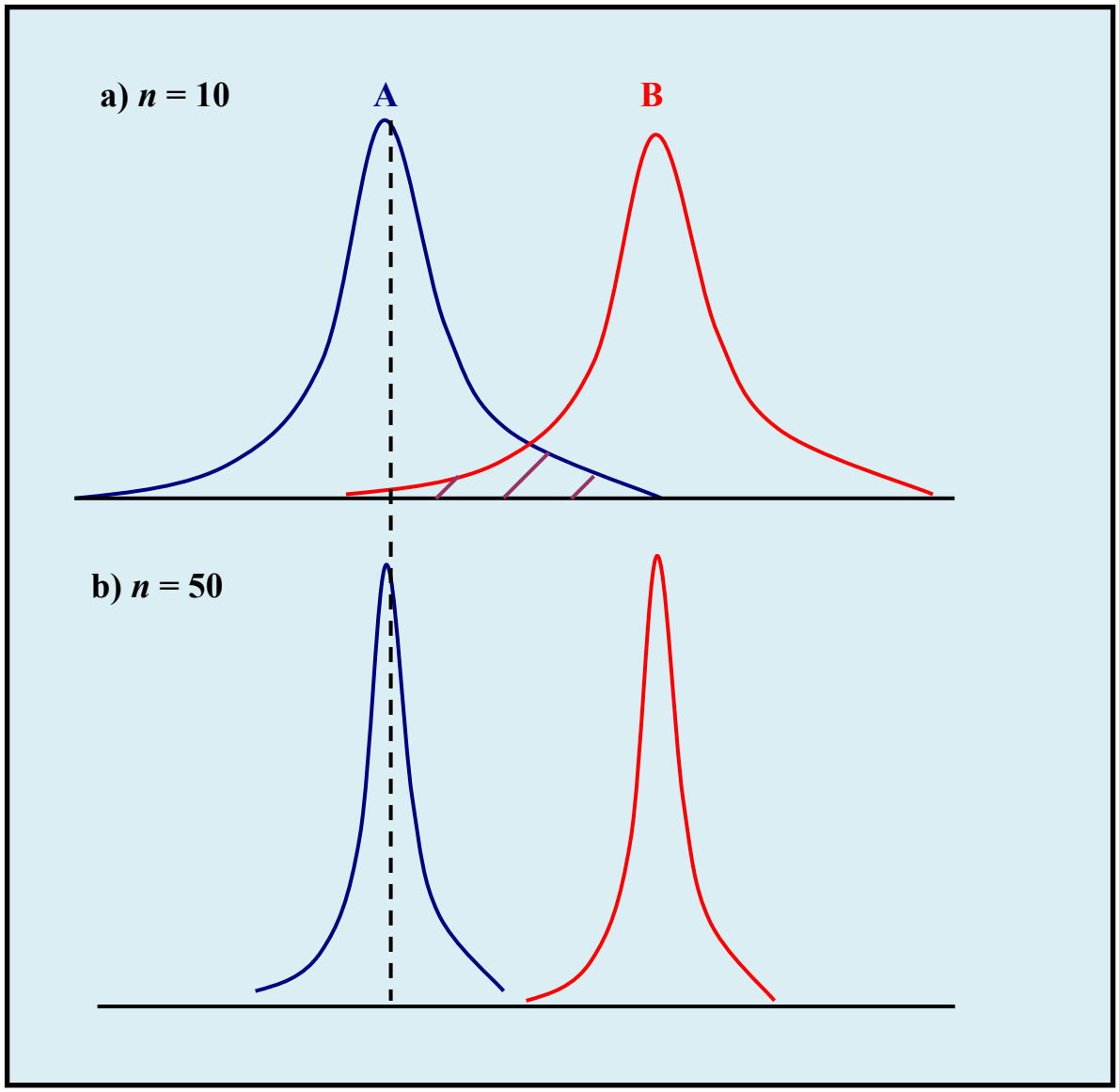


Figure 14.1 **Sampling distributions of the means for two sets of data (A and B) showing the effect of sample size (*n*) on the overlap between the two distributions: a) small sample size (*n* = 10) resulting in large standard errors of the means and overlap between both distributions, and b) larger sample size (*n* = 50) resulting in smaller standard errors of the means and virtually no overlap between the distributions. Note that the overall means of A and B remain unchanged**

Figure 14.2 shows the frequency distribution of the same data but using the difference between observed values for the means (i.e. mean $_B$ minus mean $_A$) expressed as the number of standard deviations from the overall mean (i.e. the z or t value). The blue curve shows the distribution which would be obtained if there was no true difference between the overall means of A and B; this is identical to the normalized Guassian curve (with a mean of zero and standard deviation of one) used to generate values of P for given values of z (or t). For any individual set of data the observed value of the mean is compared with a decision level (DL, which is usually z = 1.96 and equivalent to a probability of 0.05 – or 0.025 if a single sided z-test is performed) to decide whether or not to accept the *null hypothesis* (H_o) that no real difference exists and the observed data could have arisen by chance. The probability used for the decision level (the probability that the value for z could have arisen by chance) is denoted as *alpha* (α) – or $\alpha/2$ for a single sided test.

There is however a risk in relying on the null hypothesisalone. If we set alpha at 0.05 then this means that a value for z of greater than 1.96 will be used to reject the null hypothesis and we then accept that the means of the two sets of data are probably different (i.e. belong to a different frequency distribution). However, by chance, z will be greater than 1.96 (and P less than 0.05) on approximately one occasion in twenty even if no real difference exists. Therefore the chance of a false positive (i.e. accepting that there is a difference between the two means even when in fact none exists) is 1 in 20 (i.e. 0.05). Rejection of the null hypothesis when it is true is known as a *type I error*, the chance of it occurring is alpha (α). If a lower cut off point for P (i.e. α) is used then the risk of a type I error is reduced. The chance of a type I error not occurring is $(1 - \alpha)$. It is important to realize that rejection of the null hypothesis does *not* prove that the samples came from different populations – it simply means that we have failed to prove that they do *not* come from the same population.

The alternative hypothesis (H_1) is that the two sets of data belong to different populations i.e. the true difference between their means is *not* zero. Figure14.2 shows the distribution for the difference between the means plotted as z values in red. Therefore if the data belong to different populations then their values cluster around a value other than zero so that the curve shifts to the right (or to the left if there is a negative difference). The decision level cuts the H_1 curve at a different point and divides the curve into two portions. The segment to the left, denoted as *beta* (β), represents the chance of rejecting the alternative hypothesis when it is true (or not rejecting the H_0 hypothesis when it is false) and is known as a *type II error*. The difference $(1 - \beta)$ is the *probability of rejecting the null hypothesis when it is false*, and is known as the *power* of the study.

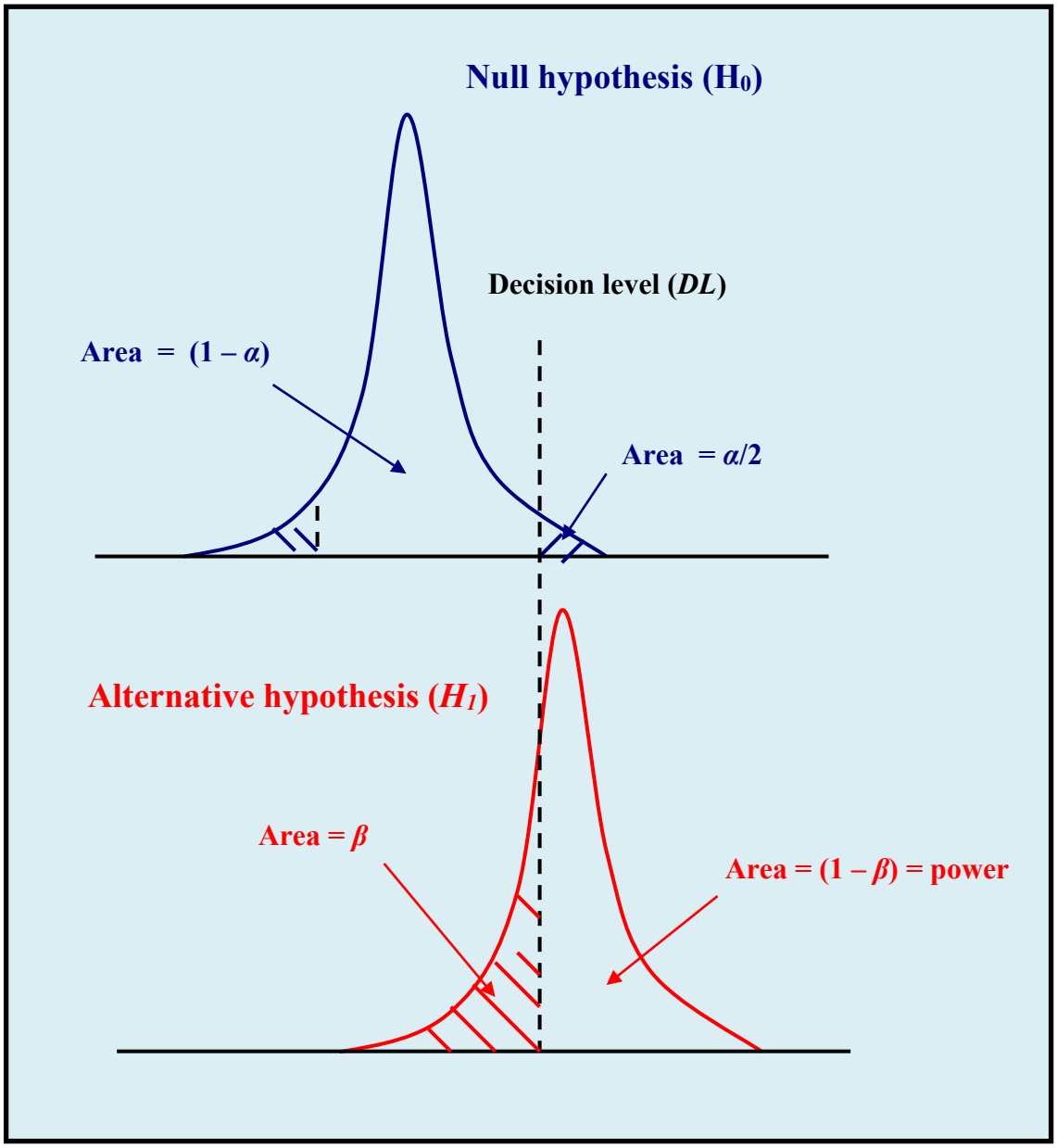


Figure 14.2 **Frequency distributions for the null (H_0) and alternative (H_1) hypotheses showing the significance of alpha (α) and beta (β)**

The relationship between type I and type II errors is shown in Fig 14.3.

	No true difference	True difference
H_0 not rejected	True result (1 - α)	Type II error (β)
H_0 rejected	Type I error (α)	True result (1 –β)

Figure 14.3 The consequences of hypothesis testing

Ideally the power should be one (or 100%) so that the null hypothesis will always be rejected whenever it is false. In practice such a high power is never achieved as there is always a small risk of a type II error. Fortunately, the power of a test can be improved by manipulating several key factors. First it is necessary to derive a relationship between these factors.

Consider two sets of overlapping data, A and B, with means m_A and m_B and a common standard deviation (s) and sample number (n). m_B is greater than m_A. The distance of the decision level (DL) from m_A is (DL - m_A) and defines the value of α. If this distance is divided by the standard error ($s/\sqrt{n}$) then the corresponding value for α expressed as standard errors from the mean (z_α) is:

$$Z_\alpha = \frac{DL - m_A}{s/\sqrt{n}} = \frac{(DL - m_A) \times \sqrt{n}}{s}$$

Similarly the distance of DL from m_B is (m_B - DL) and defines the value of β. If this distance is also divided by the standard error ($s/\sqrt{n}$) then the corresponding value for β expressed as standard errors from the mean (z_β) is:

$$z_\beta = \frac{m_B - DL}{s/\sqrt{n}} = \frac{(m_B - DL) \times \sqrt{n}}{s}$$

Adding the values for z_α and z_β gives:

$$z_\alpha + z_\beta = \frac{(DL - m_A) \times \sqrt{n}}{s} + \frac{(m_B - DL) \times \sqrt{n}}{s}$$

Which can be simplified to:

$$z_\alpha + z_\beta = \frac{(DL - m_A + m_B - DL)\sqrt{n}}{s}$$

The *DL*'s cancel giving:

$$z_\alpha + z_\beta = \frac{(m_B - m_A)\sqrt{n}}{s}$$

Substituting Δ for $(m_B - m_A)$ gives:

$$z_\alpha + z_\beta = \frac{\Delta \sqrt{n}}{s} \quad \text{..... Eq. 14.1}$$

A more useful form of this expression can be obtained by rearrangement to give a value for $\sqrt{n}$:

$$\sqrt{n} = \frac{s(z_\alpha + z_\beta)}{\Delta}$$

then squaring everything up:

$$n = [s(z_\alpha + z_\beta)/\Delta]^2 \quad \text{..... Eq. 14.2}$$

These expressions have two principal uses:

- *Evaluation of published data.* Quite often authors publish a *P* value without any discussion of the power of their study. This is particularly important when no significant difference is found since the power of the study may be too low to stand a reasonable chance of detecting any true difference which may exist. Evidence based reviews often demand some estimate of the power of published studies cited. Values for *n* and *s* and the individual means (and hence Δ) are

- usually given so it is possible to insert these into Eq 14.1 to obtain the value for ($z_\alpha + z_\beta$). The P value can be converted to z_α (by dividing it by $s/\sqrt{n}$) then z_β obtained by subtraction. z_β can be converted to β (by multiplying by $\sqrt{n}/s$) then subtracted from one to give the power. It is generally accepted that a "good" power should be at least 80%.

- *Experimental design.* Designing an experiment which has a good (at least 70-80%) power ensures a good chance of successfully detecting any effect which exists. Clearly it would only be worthwhile embarking on a study if there is a good chance of detecting a scientifically or clinically significant effect. Clinical trials, are expensive in terms of time and resources, may expose participants to some degree of risk and so clearly are irresponsible and possibly clinically unethical unless there is a reasonable chance of success. Grant awarding bodies demand evidence of power for proposed studies. At the outset levels of statistical significance (i.e. values for α and β) are set (often at 0.05 and 0.2) and it is usually possible to decide on the magnitude of clinical effect or change which would be worth detecting (i.e. a value for Δ). An estimate of the standard deviation (s) is usually available. These values can be substituted into Eq 14.2 and solved to give the sample size (n) required. It is important to remember that β always refers to a *single tailed test*.

The above principles are also applicable to other study designs but unfortunately the mathematics can become quite complex. Fortunately tables (and computer packages) are available for the common study designs so direct calculation is rarely required. Note that to simplify matters these tables often use the *effect size* (*ES*), which is the ratio of the difference between the groups to s (i.e. $ES = \Delta/s$).

Question Q 14.1

A study into the efficacy of a new cholesterol lowering drug involved measuring serum cholesterol in 30 subjects both before and after administration of the drug. Using a decision level of 5% the authors concluded that there was no effect of the drug upon serum cholesterol concentrations ($z = 0.97$, $P > 0.1$). The mean initial serum cholesterol concentration was 7.0 mmol/L ($SD = 2.0$ mmol/L). After 4 weeks of treatment with the drug the mean serum cholesterol was 6.5 mmol/L ($SD = 2.0$ mmol/L). Calculate a) the power of their study, and b) the sample size needed to achieve a power of 90%.

Answer Q 14(1)

a) The power of the study can be calculated using Eq. 14.1:

$$z_\alpha + z_\beta = \frac{\Delta \sqrt{n}}{s}$$

Δ = difference between the means for the two groups = 6.5-7.0 = - 0.5 mmol/L
n = number of subjects in the study = 30
s = standard deviation = 2.0 mmol/L

Substitution of these values for Δ, n and s into this equation allows evaluation of the sum of the z values for α and β:

$$z_\alpha + z_\beta = \frac{-0.5 \sqrt{30}}{2.0} = \frac{-0.5 \times 5.48}{2.0} = -1.37$$

Since the probability (P) used as a decision level in this study is 0.05, the corresponding z value (obtainable from tables) is 1.96. Therefore, $\alpha = 0.05$ and z_α = - 1.96 (since we are using the negative part of the distribution) allowing calculation of z_β:

$$z_\beta = -1.37 - (-1.96) = 0.59$$

From tables, the value for β (i.e. proportion of total area under the curve) corresponding to $z_\beta = 0.59$ is 0.28.

Therefore power = $(1 - \beta)$ = 1 - 0.28 = 0.72 (or 72%)

Therefore the study has a 72% chance of detected a change of 0.5 mmol/L at a decision level of 5% probability.

b) The sample size required to achieve the desired power can be calculated from Eq. 14.2:

$$n = [\, s\,(z_\alpha + z_\beta)\,/\,\Delta\,]^2$$

s = standard deviation = 2.0 mmol/L
Δ = difference between the means of the two groups = 6.5 – 7.0 = - 0.5 mmol/L

The decision level used is a probability of 0.05 (corresponding to α). From tables the corresponding z score (z_α) is 1.96 standard errors.

The required power is 90% (or 0.9).

Therefore $(1 - \beta) = 0.9$

and β =I (1.0 - 0.9) = 0.1.

From tables the corresponding z value (remembering that we are dealing with a one-tailed z-test) is 1.28.

Therefore $(z_\alpha + z_\beta) = 1.96 + 1.28 = 3.24$ standard errors

Substitute $(z_\alpha + z_\beta)$, s and Δ into the above equation and solve for n:

$$n = [\,(2.0 \times 3.24)/-0.5]^2$$

$$= 12.96^2 = \mathbf{168}$$

Therefore at least 168 subjects will need to be studied in order to achieve a 90% chance of detecting a change in serum cholesterol of 0.5 mmol/L at the 5% level of probability.

FURTHER QUESTIONS

1. A study into the effect of nutritional supplements on patients with Crohn's disease involved measuring serum albumin both before and after supplementation for a four-week period. During this period the mean serum albumin level increased from 25 g/L to 30 g/L. The study involved 40 patients with a standard deviation for albumin concentration of 10 g/L. What is the power of this study to detect a 5 g/L change in serum albumin at the 5% level of probability?

2. It is proposed to set up a study to determine the effect of dietary modification on serum cholesterol. The population to be studied has a mean serum cholesterol of 7.5 mmol/L with standard deviation of 2.5 mmol/L. What number of participants need to be recruited in order to demonstrate a lowering of serum cholesterol by 10% (using alpha = 0.05 as a critical value) with a power of 90%?

Chapter 15

Miscellaneous topics

Recovery experiments

Studies on the performance of an analytical method often include recovery experiment in which a known amount of the analyte (the "*spike*") is added to a *base* material to produce a "*spiked*" sample. The ideal outcome is that all of the added material will be recovered when the sample is analysed (i.e. the recovery is 100 percent):

$$\text{Recovery (\%)} = \frac{\text{Amount recovered} \times 100\%}{\text{Amount added}}$$

In most instances the base material will already contain some of the analyte in question but this amount is not usually known with certainty. Therefore both the base material and the spiked material are analysed. The amount recovered is calculated by subtraction of the "base result" from the "spiked result" so that the above expression becomes:

$$\textbf{Recovery (\%)} = \frac{\textbf{(spiked result - base result) x 100}}{\textbf{spike added}} \quad \textbf{…….. Eq. 15.1}$$

Allowance must be made for dilution of the spike by the base material and vice versa.

Sometimes recovery is calculated in a different way which assumes that the expected result for the spiked sample is equal to the sum of the base result and the spike added:

$$\text{Recovery (\%)} = \frac{\text{Spiked result} \times 100}{\text{Base result} + \text{spike added}}$$

This approach is fundamentally flawed and should NOT be used. The expected contribution of analyte in the base sample is not the same as its measured value (unless the recovery of the method is exactly 100%). The expected result for the analyte in the base sample is not known and therefore the expected result for the spiked base cannot be calculated and the above equation does not give a valid recovery value. Only the value for the "spike" is accurately known.

Most assays in routine use have recoveries in the range 90-100% with imprecisions of the order of 5%. Imprecision of the assay can easily lead to unreliable estimates of recovery and can be minimized by:

- Spiking the base material with concentrations similar to the endogenous analyte concentration.
- Performing several replicates.

Sometimes it is useful to carry out recovery experiments with a range of spiked values and using several different base materials.

Question Q 15(1)

A plasma glucose assay involved adding 0.1 mL of sample to 2.0 mL of reagent, incubating at 37ºC for 30 minutes then measuring the absorbance at 500 nm. Using 0.1 mL of water as sample the absorbance reading was 0.080, using 0.1 mL of a 10 mmol/L glucose standard it was 0.320 and using a patient's plasma sample 0.200.

0.1 mL of a 50 mmol/L glucose solution was mixed with 0.9 mL of the same plasma sample, then 0.1 mL of the mixture taken through the assay. The absorbance was 0.300. Calculate the recovery of the method.

Answer Q 15(1)

First calculate the concentration of glucose in the un-spiked plasma sample:

Assuming that absorbance is directly proportional to glucose concentration:

$$\frac{\text{Plasma glucose (mmol/L)}}{\text{Glucose standard (mmol/L)}} = \frac{\text{Plasma absorbance}}{\text{Absorbance of standard}}$$

Rearranging:

$$\text{Plasma glucose (mmol/L)} = \frac{\text{Plasma absorbance x glucose standard (mmol/L)}}{\text{Absorbance of standard}}$$

Subtract the blank (water used as sample) absorbance from both standard and plasma readings and substitute into the above equation:

$$\text{Plasma glucose} = \frac{(0.200 - 0.080) \times 10}{(0.320 - 0.080)} = \frac{0.12 \times 10}{0.24} = 5 \text{ mmol/L}$$

It is not necessary to calculate absolute amounts of glucose in the samples in order to calculate recovery; concentrations can be used:

$$\text{Recovery (\%)} = \frac{\text{Recovered concentration x 100}}{\text{Concentration added}}$$

However, by mixing the diluted glucose solution with the plasma sample the base concentration and the spiked concentration has been diluted and must first be calculated:

$$\text{Glucose from plasma} = \frac{\text{Plasma glucose (mmol/L) x Plasma volume (mL)}}{[\text{Volume of plasma (mL)} + \text{volume of glucose solution (mL)}]}$$

$$= \frac{5 \times 0.9}{(0.9 + 0.1)} = 4.5 \text{ mmol/L}$$

$$\text{Glucose added} = \frac{\text{Concn of glucose solution (mmol/L) x Volume of glucose added (mL)}}{\text{[Vol of glucose added (mL) + Volume of plasma (mL)]}}$$

$$= \frac{50 \times 0.1}{(0.9 + 0.1)} = 5 \text{ mmol/L}$$

Next calculate the measured glucose in the spiked plasma sample – remembering to subtract the reagent blank:

$$\text{Spiked sample glucose concentration} = \frac{\text{Spiked sample absorbance x Standard concn}}{\text{Absorbance of standard}}$$

$$= \frac{(0.300 - 0.080) \times 10}{(0.320 - 0.080)}$$

$$= \frac{0.22 \times 10}{0.24} = 9.17 \text{ mmol/L}$$

Calculate the recovery from the measured glucose concentration in the spiked plasma (9.17 mmol/L), the base plasma sample (4.5 mmol/L) and the glucose added (5 mmol/L):

$$\text{Recovery (\%)} = \frac{\text{[Glucose in spiked plasma (mmol/L) - glucose in base plasma (mmol/L)] x 100}}{\text{Glucose added to base plasma (mmol/L)}}$$

$$= \frac{(9.17 - 4.5) \times 100}{5} = \frac{4.67 \times 100}{5} = \mathbf{93\%} \text{ (2 sig figs)}$$

Elimination of a tumour marker

Frequently tumour markers are measured following surgical resection of the tumour to provide evidence that removal of the tumour is complete. If the tumour has been completely removed then the concentration of the tumour marker should reflect the value expected from the natural decay of the compound present at the time of surgery. If the concentration has not fallen to this level then it is likely that the marker is still being produced by residual tumour tissue (or from secondaries).

Tumour markers are usually cleared exponentially as are most drugs (see Chapter 7). If

Cp_t = concentration of tumour marker at time t
Cp_0 = initial concentration of tumour marker
k_d = elimination rate concentration

then as was shown in Chapter 7, a linear expression can be derived relating these variables:

$$\ln Cp_t = \ln Cp_0 - k_d.t \quad \text{..... Eq. 7.7}$$

Provided the rate constant k_d is known then the concentration at any given time can be calculated or the time taken to reach a specified concentration estimated. The latter may be particularly useful in predicting the time when the concentration of tumour marker should be below the upper reference limit.

There are ways of manipulating this equation to give an expression that is simpler to apply in common situations. Subtraction of one logarithm from another is the same thing as calculating the logarithm of their ratios:

Therefore $\ln Cp_t - \ln Cp_0 = \ln (Cp_t/Cp_0) = \ln CR$

Where CR is the ratio of the concentration at time t (CP_t) to the initial concentration (Cp_0). Eq. 7.7 can then be re-written as:

$$\ln CR = - k_d.t \quad \text{..... Eq. 15.2}$$

Further simplification is possible. In Chapter 7 it was shown that the elimination rate constant is related to the half-life ($t_{1/2}$):

$$t_{1/2} = \frac{0.693}{k_d} \quad \text{..... Eq. 7.9}$$

Substitution of $k_d = 0.693/t_{1/2}$ into expression Eq 15.2 gives:

$$\ln CR = - \frac{0.693.\, t}{t_{1/2}} \quad \text{..... Eq. 15.3}$$

If the time (t) is expressed as the number of half lives (N) so that $N = t/t_{½}$ then an even simpler expression is produced:

$$\ln CR = - 0.693\, N \quad \text{..... Eq. 15.4}$$

If logarithms to the base 10 are used then 0.693 is divided by 2.303 (since $\ln CR = 2.303 \log_{10} CR$) and the expression becomes:

$$\log_{10} CR = - 0.30\, N \quad \text{..... Eq. 15.65}$$

Using this expression it is quite simple to determine the number of half-lives required for a given change in concentration ratio or if the absolute time period is known then the half-life of the tumour marker can be determined.

Question Q 15(2)

A tumour marker X is used to guide a decision on chemotherapy after the resection of the main tumour mass. The concentration decays exponentially. If the half-life of the tumour marker is less than 75 hours, then this is indicative of tumour clearance and chemotherapy is withheld. If the half-life is greater than this, it indicates that residual disease is present and chemotherapy is indicated. The precision of the assay is such that measurements can be safely made at a precisely timed interval of more than 36 hours from two or more days after surgery.

The level of X at 50 hours post surgery is 1756 ng/L and at 94 hours it is 1050 ng/L. Calculate the half-life and indicate whether you can say with confidence whether chemotherapy needs to be given.

Answer Q 15(2)

There are two ways of solving this problem:

1. *Using Eq 7.7:*

$$\ln CP_t = \ln Cp0 - k_d.t$$

The 1st sample (collected at 50 h) can be regarded as the initial sample. Therefore:

CP_t = tumour marker concentration at 94 h = 1050 ng/L
Cp_0 = tumour marker concentration at 50 h = 1756 ng/L
t = time interval between the two samples = 94 - 50 = 44 h

Substituting these values into Eq. 7.7 allows calculation of k_d:

$$\ln 1050 = \ln 1756 - 44\ k_d$$

$$6.96 = 7.47 - 44\ k_d$$

$$44\ k_d = 7.47 - 6.96 = 0.51$$

$$k_d = \frac{0.51}{44} = 0.0116\ h^{-1}$$

Use Eq. 7.9 to covert the elimination rate constant (k_d) to the half life ($t_{½}$):

$$t_{½} = \frac{0.693}{k_d} = \frac{0.693}{0.0116} = \mathbf{60\ h}\ (2\ sig\ figs)$$

Since the half-life is less than 75 h, the time interval is 44 h and the 1st sample was taken at least 48 h after tumour removal we can conclude that chemotherapy can be withheld.

2. *Using Eq 15.5*

$$\log_{10} CR = - 0.30\, N$$

$$CR = \text{concentration ratio} = \frac{Cp_t}{Cp_o} = \frac{1050}{1756} = 0.598$$

Substituting $CR = 0.598$ gives: $\log_{10} 0.598 = - 0.30\, N$

Rearranging and solving for N:

$$N = - \frac{\log_{10} 0.598}{0.30} = - \frac{(- 0.223)}{0.30} = 0.744 \text{ half lives}$$

Thus the concentration fell from 1756 ng/L to 1050 ng/L in 0.744 half-lives OR 44 h. This information can be used to calculate the half-life:

$N = t/t_{½}$ so that $t_{½} = t/N$.

Substitute $t =$ 44 h and $N = 0.744$ and solve for $t_{½}$:

$$t_{½} = \frac{44}{0.744} = \textbf{59 days} \text{ (2 sig figs)}$$

Radioactive decay

Radioactive decay follows first order kinetics and therefore obeys the same mathematical laws as the clearance of the drug (or elimination of a tumour marker discussed in the preceding section). Equations Eq. 7.7 and 15.5 can be used if units of radioactivity are substituted for concentration.

Question Q 15(3)

If the half life of a radionucleotide is 20 hours at the end of how many complete days will the activity have fallen to less than 2% of the initial value?

Answer Q 15(3)

Using equation Eq. 15.5:

$$\log_{10} CR = - 0.30\, N$$

Rearrangement gives:

$$N = \frac{- \log_{10} CR}{0.30}$$

Since the activity falls from 100% to 2% of the initial value, then these figures can be treated as concentrations Cp_0 and Cp_t respectively:

$$CR = \frac{Cp_t}{Cp_0} = \frac{2}{100} = 0.02$$

Substituting $CR = 0.02$ permits calculation of N:

$$N = \frac{- \log_{10} 0.02}{0.30} = \frac{- (-1.70)}{0.30} = \frac{1.70}{0.30} = 5.67$$

Therefore the activity reached 2% of the initial value after 5.67 half lives have elapsed. Multiplication of 5.64 by the half-life (20 h) gives the total time period (t) in hours.

$$t = 5.67 \times 20 = 113.4$$

Division by 24 gives the number of days:

$$t = \frac{113.4}{24} = 4.7 \text{ days}$$

Therefore **5** complete days must elapse before the activity falls below 2% of the initial value.

Exponential growth

The mathematics of exponential growth are the same as exponential decay or elimination except that concentration increases with time. The elimination constant (k_d) is replaced by the specific growth rate (k) and the term $-k_d.t$ becomes positive ($+k.t$). As a consequence Eq. 7.7 and Eq. 15.5 become:

$$\ln Cp_t = \ln Cp_0 + k.t \quad \text{..... Eq. 15.6}$$

$$\log_{10} CR = 0.30\,N \quad \text{..... Eq. 15.7}$$

The half-life ($t_{½}$) is replaced by doubling time (t_d) so that Eq. 7.9 can be re-written:

$$t_d = \frac{0.693}{k} \quad \text{..... Eq. 15.8}$$

Question Q 15(4)

A woman had a beta hCG concentration measured at 265 IU/L and 11 days later, following some abdominal pain, it was 820 IU/L. Assuming hCG rises exponentially in early pregnancy, what has been the doubling time over this period? What is the significance of the result you obtain?

Answer Q 15(4)

This problem can be solved by direct application of Eq. 15.7:

$$\log_{10} CR = 0.30\, N$$

Rearranging gives: $$N = \frac{\log_{10} CR}{0.30}$$

CR can be calculated taking the initial concentration (Cp_0) as 265 IU/L, and the Cp_t as 820 IU/L:

$$CR = \frac{Cp_t}{Cp_0} = \frac{820}{265} = 3.095$$

Substituting $CR = 3.095$ and solving for N:

$$N = \frac{\log_{10} CR}{0.30} = \frac{\log_{10} 3.095}{0.30} = \frac{0.491}{0.30} = 1.64$$

The doubling time (t_d) can be calculated from N (1.63) and the time period between measurements (t = 11 days) using Eq. 15.8:

$$t_d = \frac{t}{N} = \frac{11}{1.63} = \textbf{6.7 days}$$

The normal doubling time for hCG during early pregnancy is approximately 2 days. Therefore this result is consistent with ectopic pregnancy.

Urinary nitrogen excretion and nitrogen balance

Measurements of urinary nitrogen excretion are often used to provide an estimate of nitrogen balance in patients on pareneteral nutrition. Nitrogen intake can be calculated from the amount of feeding solutions given but measurements of total urinary nitrogen are cumbersome and not widely available. Instead nitrogen excretion is estimated from the 24 urinary urea excretion. The underlying assumption is that all amino acid nitrogen is converted to urea, is not excreted by any other route (i.e. the gut, sweat or fistulae) and that there is no other significant nitrogen loss in the urine (i.e. amino acids, ammonia etc).

Urea has the formula $CO(NH_2)_2$ so that each molecule contains two atoms of nitrogen. Therefore each mmol of urea contains 2 mmol of nitrogen and since the atomic weight of nitrogen is 14, this constitutes 2 x 14 = 28 mg of nitrogen. Division by 1000 converts this figure to g of nitrogen:

$$\textbf{Nitrogen excretion (g/24 h)} = \frac{\textbf{Urea excretion (mmol/24 h) x 28}}{\textbf{1000}} \quad \textbf{……. Eq. 15.9}$$

It has been advocated that a figure of 20% should be added to nitrogen excretion to allow for other urinary losses and a further 2 g/day added to allow for losses by other routes.

Subtraction of this figure from the nitrogen intake gives the nitrogen balance:

$$\textbf{Nitrogen balance (g/24 h)} = \textbf{Nitrogen intake (g/24 h)} - \textbf{Nitrogen excretion (g/24 h)} \quad \textbf{…….. Eq. 15.10}$$

Question Q 15(5)

A patient receiving total parenteral nutrition is receiving 12 g nitrogen/24 h as amino acids. Urinary urea excretion is 600 mmol/24 h. Indicating what assumptions you make, calculate whether she is in positive or negative nitrogen balance.

Answer Q15(5)

Use Eq. 15.9 to calculate nitrogen excretion from urea excretion:

$$\text{Urinary nitrogen excretion (g/24 h)} = \frac{\text{Urinary urea excretion (mmol/24 h)} \times 28}{1000}$$

$$= \frac{600 \times 28}{1000} = 16.8 \text{ g/24 h}$$

Use Eq. 15.10 to calculate nitrogen balance:

$$\text{Nitrogen balance (g/24 h)} = \text{Nitrogen intake (g/24 h)} - \text{Nitrogen excretion (g/24 h)}$$

$$= 12 - 16.8 = \mathbf{-4.8 \text{ g/24 h}}$$

Correcting for other urinary losses (+20%) and other routes of nitrogen excretion (+2 g) gives a revised value for nitrogen excretion:

$$\text{Corrected nitrogen excretion} = \frac{(16.8 \times 120)}{100} + 2 = 22.2 \text{ g/24 h}$$

So that the corrected nitrogen balance becomes:

$$\text{Corrected nitrogen balance} = 12 - 22.2 = \mathbf{-10.2 \text{ g/24 h}}$$

Tritratable gastric acidicity

Acid secretion studies are less frequently performed nowadays, the main remaining application is in the follow-up of a raised plasma gastrin result. The documentation of a raised basal acid output (BAO) in gastric juice provides strong evidence that a high plasma gastrin concentration is caused by Zollinger-Ellison syndrome. After an overnight fast gastric juice is collected over a timed period (usually 30 min), its volume measured and an aliquot titrated with standardised sodium hydroxide:

HCl	+	NaOH	→	NaCl	+	H_20
In gastric fluid		Known concentration				

There has been considerable debate over the years as to which pH to use as the endpoint since secreted hydrochloric acid may be buffered by gastric proteins.

At the equivalence point the total amount of acid in the gastric fluid aliquot is equal to the amount of sodium hydroxide added. The total amount of acid or alkali is equal to the volume used multiplied by its concentration. Therefore, we can write:

$$M_1 \times V_1 = M_2 \times V_2 \quad \text{..... Eq. 15.11}$$

where M_1 = molar concentration of hydrochloric acid in the gastric fluid
V_1 = volume of aliquot of gastric fluid used in the titration
M_2 = molar concentration of sodium hydroxide
V_2 = titre of sodium hydroxide solution

Eq. 15.11 can be rearranged to determine M_1 which can be converted to the total acid secreted and finally the rate of acid secretion.

Question Q 15(6)

During a test of gastric acid secretion, 28mL of gastric juice was aspirated over a single 30 min period from a fasting patient before the administration of pentagastrin. The volume of 0.1 M sodium hydroxide solution required to titrate a 5 mL aliquot of the gastric juice to pH 7.4 was 14 mL. Calculate the basal acid secretion rate in mmol/h.

Answer Q 15(6)

Since the answer is required in mmol/h it is easiest to work in mmol from the outset. Use Eq. 15.11:

$$M_1 \times V_1 = M_2 \times V_2$$

M_1 = millimolar concentration of acid in gastric fluid = ? mmol/L
V_1 = volume of gastric fluid aliquot used in titration = 5 mL
M_2 = millimolar concentration of sodium hydroxide = 0.1 x 1000 = 100 mmol/L
V_2 = titre of sodium hydroxide solution = 14 mL

Substitute for V_1, M_2 and V_2 in Eq. 15.11 and solve for M_1:

$$M_1 \times 5 = 100 \times 14$$

$$M_1 = \frac{100 \times 14}{5} = 280 \text{ mmol/L}$$

Multiply the acid concentration (280 mmol/L) by the gastric fluid volume in litres (28 mL divided by 1000) to give the total acid output in millimols:

$$\text{Total acid output} = \frac{280 \times 28}{1000} = 7.84 \text{ mmol}$$

Multiply by 2 (since gastric fluid was collected over 30 min) to obtain the secretion rate in mmol/h:

Secretion rate = 7.84 x 2 = **16 mmol/h** (2 sig figs)

Internal standardisation

Chromatographic techniques are often subject to variability (e.g. due to instability of the detector) or to unpredictable losses (due to the preliminary sample preparation which may involve extraction and/or derivatization). Addition of an *internal standard*, which is a compound chemically related to the analyte with similar properties but resolvable by chromatography, can be used to minimize these problems. Instead of plotting the peak height or area of the standard against standard concentration, the ratio of peak height (or area) of the standard to that of the internal standard is used. An example is shown in Figure 15.1 in which both methods of plotting the standard curve are used. The same amount of internal standard is also added to each of the unknown samples, the peak height or (area) ratio similarly calculated and the results obtained by interpolation with the standard curve.

The underlying assumption of this approach is that the internal standard behaves exactly as the analyte being measured i.e. the ratio of analyte to internal standard remains constant throughout the various stages of the analytical process.

Question Q 15(7)

An HPLC method for estimation of plasma phenylalanine uses N-methyl 1-phenylalanine (NMP) as internal standard. 200 µL of NMP has been added to 200 µL aliquots of standard or sample prior to analysis. The following peak areas were obtained:

Sample	**Peak area**	
	NMP	**Phenylalanine**
Standard (500 µmol/L)	20,000	81,000
Patient	18,000	120,000

Calculate the phenylalanine concentration in the patient's sample.

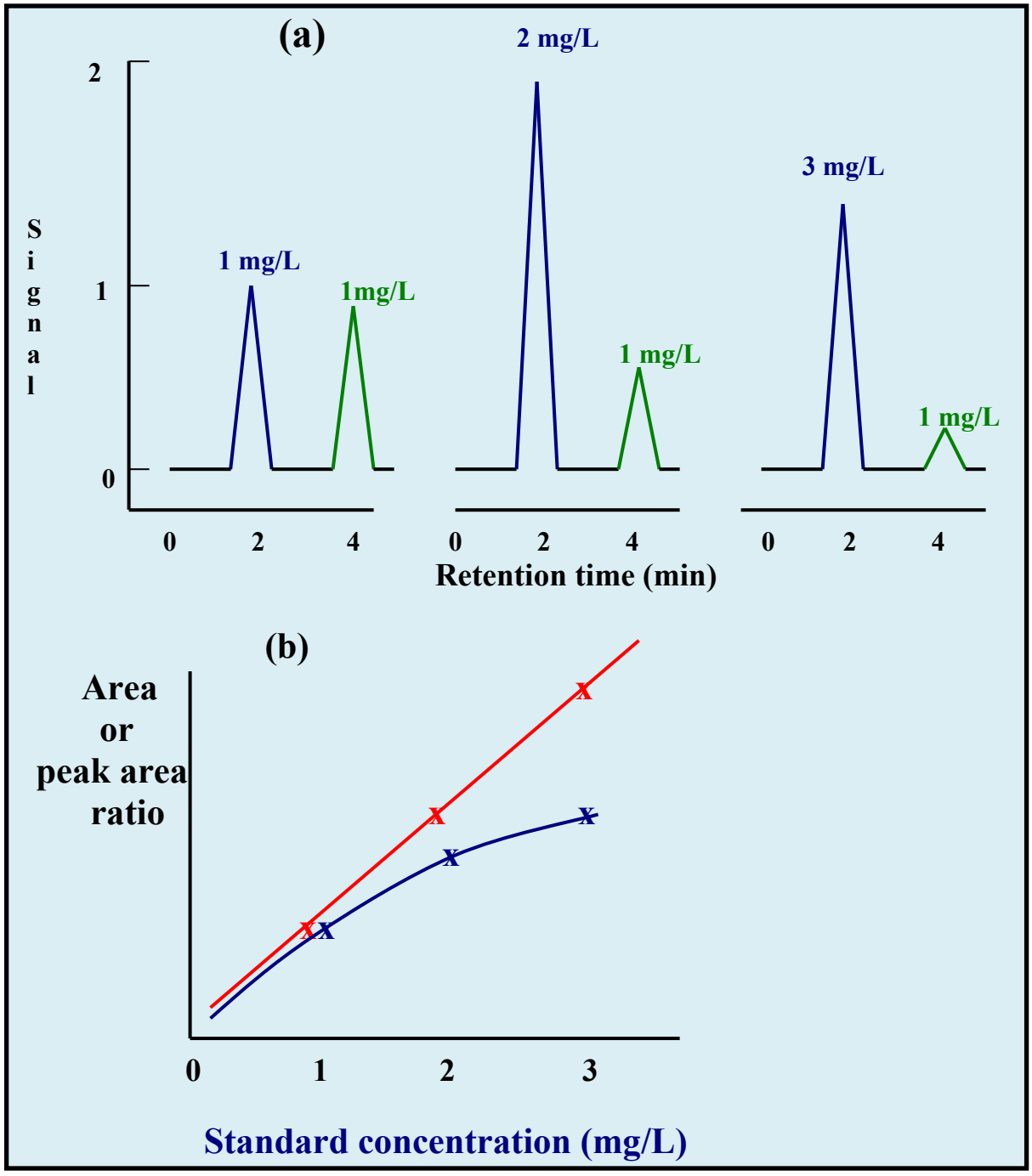


Figure 15.1 a) Chromatographs of standards at three levels, spiked with equal amounts of internal standard (IS), b) Standard curves of standard peak area and standard:internal standard peak area ratios

Answer Q 15(7)

First calculate the peak area ratio (PAR) of the phenylalanine peak to that of the internal standard (NMP) for the standard and patient:

$$\text{PAR (Standard)} = \frac{81,000}{20,000} = 4.05$$

$$\text{PAR (Patient)} = \frac{120,000}{18,000} = 6.67$$

Assuming that PAR is proportional to concentration

$$\frac{\text{PAR (standard)}}{\text{Concn (standard)}} = \frac{\text{PAR (unknown)}}{\text{Concn (unknown)}}$$

Which can be rearranged to give:

$$\text{Concn (unknown) (µmol/L)} = \frac{\text{PAR (unknown)} \times \text{Standard concn (500 µmol/L)}}{\text{PAR (Standard)}}$$

$$\text{Patient phenylalanine concn} = \frac{6.67 \times 500}{4.05} = \mathbf{823}\ \mu\text{mol/L}$$

Population genetics – the Hardy-Weinberg equilibrium

Consider an *ideal* population in which there is a gene locus with two alleles *A* and *a* with gene frequencies *p* and *q* i.e:

p = frequency of the dominant allele (*A*)
q = frequency of the recessive allele (*a*)

Mating results in the random combination of a male gamete (*A* or *a*) with a female gamete (*A* or *a*). The possible results, with their frequencies, can be obtained by constructing a *Punnett's square* (Fig 15.2):

		Male gametes	
		A **(p)**	***a*** **(*q*)**
Female gametes	***A* (*p*)**	***AA*** **(p^2)**	***Aa*** **(*pq*)**
	***a* (*q*)**	***Aa*** **(*pq*)**	***aa*** **(q^2)**

Figure 15.2 Punnetts square showing the allele frequencies and resulting genotype frequencies for a two allele system (*A* and *a* with frequencies *p* and *q*) in the first generation

Therefore the recombination of *A* and *a* gametes from both parents results in three genotypes – *AA*, *Aa* and *aa* with frequencies p^2, $2pq$ and q^2. Since the frequencies of all genotypes must add up to 1 the following useful expressions can be written:

For a pair of alleles: $p + q = 1$ ..... **Eq. 15.12**

For genotypes: $p^2 + 2pq + q^2 = 1$ ..... **Eq. 15.13**

Mathematically, Eq. 15.13 is the expansion of $(p + q)^2$.

Subsequent mating of this population to produce a second generation can only produce offspring with genotypes *AA*, *Aa* or *aa*. The frequencies can be calculated from the frequencies of the possible mating combinations:

For example *AA* can arise from the combinations *AA* x *AA*, *AA* x *Aa* and *Aa* x *Aa*. The frequencies can be derived from the frequencies derived after the first cross as follows:

AA x *AA* has only one possible outcome (*AA*) , therefore frequency of *AA* p^2 x p^2 = p^4

AA x *Aa* can give rise to *AA* in two ways: *A* from the male or *A* from the female. The total initial frequency of *Aa* is $2pq$ (one pq from males and one pq from females). Therefore the frequency of *AA* in this cross is p^2 x $2pq$ = $2p^3q$

Aa x *Aa.* The initial frequency of each of these is $2pq$, so that the frequency of the *A* gamete from each parent is pq . Therefore frequency of *AA* in the second cross = pq x pq = p^2q^2

Adding the frequency of *Aa* offspring gives $p^4 + 2p^3q + p^2q^2$, which can also be written as factors with p^2 outside the bracket: $p^2 (p^2 + 2pq + q^2)$. Since we know that $(p^2 + 2pq + q^2) = 1$ (Eq. 15.14), the frequency of *AA* remains at p^2.

Repeating this process for all the possible matings in the second generation gives the data in Fig 15.3. The frequency of the various genotypes after this second mating remains unchanged. This calculation could be repeated for further generations but the frequencies of genotypes *AA*, *Aa* and *aa* would always remain at p^2, $2pq$ and q^2.

Mating type	Frequency	Frequency of offspring		
		AA	*Aa*	*aa*
AA x *AA*	p^4	p^4	-	-
AA x *Aa*	$4p^3q$	$2p^3q$	$2p^3q$	-
Aa x *Aa*	$4p^2q^2$	p^2q^2	$2p^2q^2$	p^2q^2
AA x *aa*	$2p^2q^2$	-	$2p^2q^2$	-
Aa x *aa*	$4pq^3$	-	$2pq^3$	$2pq^3$
aa x *aa*	q^4	-	-	q^4
Total		$p^2(p^2+2pq+q^2)$	$2pq(p^2+2pq+q^2)$	$q^2(p^2+2pq+q^2)$
Relative frequency		p^2	$2pq$	q^2

Figure 15.3 Frequency of various types of offspring due to different matings in the second generation

This principle was discovered by Godfrey Hardy and Wilhelm Weinverg and is known as the *Hardy-Weinberg equilibrium*:

> *"Gene frequencies and genotype ratios in a randomly breeding population remain constant from generation to generation".*

For this equilibrium to hold true seven conditions must be met:

- mutation is not occurring
- natural selection is not occurring
- the population is infinitely large
- all members of the population breed
- all mating is totally random
- everyone produces the same number of offspring
- there is no migration in or out of the population

Identity of p and q for a population depends on the phenotypic expression of the alleles:

1. In-complete dominance i.e. homozygosity for the dominant gene is required in order for the phenotype to be expressed. In this instance the proportion of the population expressing the genotype (*AA*) is the same as p^2. p is then the square root of this value. Subtraction of p from 1 gives q (Eq. 15.12).

2. Complete dominance i.e. both homozygotes (*AA*) and heterozygotes (*Aa*) express the phenotype. Therefore, the proportion of individuals *without* this phenotype gives the value of q^2. q is the square root of this value. Subtraction of q from 1 gives p.

3. Autosomal recessive i.e. only those individuals homozygous for the recessive gene (*aa*) express the phenotype. Therefore the proportion of individuals *with* this phenotype represent the q^2 frequency. q is the square root of this value. Subtraction of q from 1 gives p.

Question Q 15(8)

Phenylketonuria is inherited in an automal recessive manner. If the prevalence of the disorder is 10 in 10,000 estimate the percentage of heterozygous carriers in the population.

Answer Q 15(8)

For an autosomal recessive mode of inheritance the frequency of the phenylketonuria phenotype represents the autosmal recessive genotype i.e. q^2

$$q^2 = 1 \text{ in } 10{,}000 = \frac{1}{10{,}000} = 0.0001$$

$$q = \sqrt{0.0001} = 0.01$$

Using Eq. 15.12: $p + q = 1$

$$p = 1 - q = 1 - 0.01 = 0.99$$

The carriers are heterozygous with a frequency of 2pq. Therefore:

Frequency of carriers = 2 x 0.99 x 0.01 = 0.0198

Multiply by 100 to convert to percentage = 0.0198 x 100 = **1.98 (approx 2%)**

When data can be broken down into a set of compartments (i.e. grouped frequencies), for each of which there is an observed number (O) and an expected (E) number of individuals, the goodness-of fit- (X^2) is given by:

$$X^2 = \sum \frac{(O - E)^2}{E} \quad \text{..... Eq. 15.14}$$

Care must be taken to use the correct number of degrees of freedom. For a contingency table of r rows and c columns there are $(r - 1)(c - 1)$ degrees of freedom. If the data is only grouped into k classes then there are $(k - 1)$ degrees of freedom. The probability of obtaining any value of X^2 can be obtained by looking up its P value in tables of X^2. In general, every additional restraint removes one degree of freedom.

The X^2 test is extremely useful in genetics to assess whether the distribution of genotypes fits the predicted pattern. For example if the observed and expected (i.e. predicted) frequencies were:

Genotype	Observed	Expected
AA	80	100
Aa	240	200
Aa	80	100

$$X^2 = \sum \frac{(O - E)^2}{E} = \frac{(80 - 100)^2}{100} + \frac{(240 - 200)^2}{200} + \frac{(80 - 100)^2}{100}$$

$$= 16.0 \quad (2 \text{ degrees of freedom})$$

From tables of X^2, the P value for a X^2 value of 16 with 2 degrees of freedom is less than 0.001. Therefore the observed frequencies do not fit with the predicted model.

It is very important to appreciate that the X^2 test should ONLY be used for grouped frequencies. This is a frequently abused test. In spite of the fact that Eq. 15.14 uses the terms "observed" and "expected" results, it should NOT be used evaluate data from method comparison studies.

Question Q 15(9)

The prevalence of a metabolic disease which is inherited in an autosomal recessive manner due to a single allele is 1 in 10,000. A survey identified 1% of the population as asymptomatic carriers. Are these data consistent with the population in a Hardy-Weinberg equilibrium?

Answer Q 15(9)

Let the dominant gene for the disorder be *A* and the recessive gene *a*. As the inheritance of the disease is autosomal recessive only the homozygous recessive genotype (*aa*) expresses the disease.

The incidence of the receive disorder (*aa*) = 1 in 10,000 = $\frac{1}{10{,}000}$ = 0.0001

The incidence of carriers (*Aa*) is 1 % = $\frac{1}{100}$ = 0.01

Since the total must equal 1, the incidence of the homozygous dominant genotype (which does not express disease nor have carrier status) can be calculated:

Incidence of homozygous dominant (*AA*) = 1 - (0.0001 + 0.01)

= 1 - 0.0101 = 0.9899

To summarise the observed frequencies of the three genotypes:

Genotype	*AA*	*Aa*	*aa*
Observed frequency	0.9899	0.01	0.0001

To calculate the expected frequency if Hardy-Weinberg equilibrium is present, determine values for *p* and *q* then calculate the expected frequencies for the three genotypes.

Frequency of affected individuals (*aa*) = 0.0001 = q^2

Therefore $q = \sqrt{q^2} = \sqrt{0.0001} = 0.01$

Since $p + q = 1$, $p = 1 - q$

Therefore $p = 1 - 0.01 = 0.99$

Using these values of p and q the expected frequencies can be calculated as follows:

Frequency of AA = p^2 = 0.99^2 = 0.9801

Frequency of Aa = $2\,p\,q$ = 2 x 0.99 x 0.01 = 0.0198

Frequency of aa = q^2 = 0.01^2 = 0.0001

If all the data is tabulated then X^2 can be calculated at the same time:

Genotype	Frequency		$(O - E)$	$(O - E)^2$	$(O - E)^2/E$
	Observed	***E***xpected			
AA	0.9899	0.9801	0.0098	0.00009604	0.00009799
Aa	0.01	0.0198	- 0.0098	0.00009604	0.00485051
Aa	0.0001	0.0001	0.0000	0.00000000	0.00000000
Total	1.0000	1.0000	0.0000	0.00019208	0.00497509

X^2 is the sum of all the values in the final column = 0.0050 (2 sig figs)

Normally the degrees of freedom would be 3 - 1 = 2. However, since one of the observations (frequency of the disease) was used to estimate the expected values, a further degree of freedom is lost leaving only one.

From tables, the value of P for X^2 = 0.0050 is somewhere between 0.05 and 0.95. Therefore there is no significant difference between the observed and expected values so that the data fit with the Hardy-Weinberg equilibrium.

Competitive binding assays

Many assays in laboratory medicine utilise the property of the analyte in question to bind specifically and reversibly to a protein. More often than not this protein is an *antibody* and so these techniques are known as *immunoassays*. The antibody (Ab) binds to the analyte in question, the *antigen* (Ag) to form the *antibody-antigen complex* (AbAg) according to the equilibrium:

$$Ab + Ag \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} AbAg$$

$$K = \frac{k_1}{k_{-1}} = \frac{[AbAg]}{[Ab]\,[Ag]}$$

where

k_1 = the rate constant for the forward reaction
k_{-1} = the rate constant of the reverse reaction
K = the *binding constant* for the overall reaction

If antigen is labelled in some way (Ag*) and if the label does not interfere with binding then both the labelled and unlabelled antigen compete for the binding sites according to the scheme:

$$\begin{array}{ccccc} Ab & + & Ag^* & \rightleftharpoons & AbAg^* \\ + & & & & \\ Ab & & & & \\ \downarrow\uparrow & & & & \\ AbAg & & & & \end{array}$$

This is analogous to the scheme for competitive inhibition of an enzyme except that the antibody-antigen complexes do not break down to produce reaction products.

Since the labelled and unlabelled antigen compete for the same binding sites it follows that if the biological sample containing unlabelled antigen and a reagent with labelled antigen are both added to a solution containing the antibody then the proportion of AbAg* and AbAg formed will reflect the proportion of Ag* and Ag in the reaction mixture. In fact the concentration of AbAg* will be inversely proportional to the concentration of Ag in the sample. This is the basic principle of all competitive binding assays. A plethora of techniques have been developed which differ in the nature of the label used, whether separation of bound from free label is necessary (and the technique used to achieve this), the nature of the binding protein etc. Additionally non-competitive assays have been developed.

The simplest type of immunoassay is probably the radioimmunoassay in which the analyte competes for antibody binding sites with an antigen to which a radioisotope label has been attached. At equilibrium we are left with a reaction mixture which contains free analyte (Ag), bound analyte (AbAg), free label (Ag*), bound label (AbAg*) and free antibody (Ab). In order to be able to measure the radioactivity due to bound label alone it is necessary to separate bound from free label i.e. AbAg* from Ag*. This is usually achieved by precipitation (e.g. using a second antibody or polyethylene glycol). The radioactivity in the precipitate is a measure of AbAg*. The higher the concentration of antigen in the sample (Ag) the lower the radioactivity in the precipitate. Unfortunately there is no simple mathematical relationship between the analyte concentration in the sample (Ag) and the radioactivity in the precipitate. Numerous complex mathematical procedures have been devised to enable calculation of analyte concentration from the count rate which are beyond the scope of this book. However, in many cases it is possible to obtain a reasonable linear relationship over a limited working range for the assay by plotting the proportion of label bound (usually expressed as a percentage) against the logarithm of analyte concentration (Fig 15.4). It is usual to carry out all measurements in duplicate to minimize imprecision.

In order to express the count rate (B) as a percentage of the maximum binding (B_0) the total binding (TB) is determined using a standard containing zero concentration of analyte so that there is no competition for binding sites. In practice there is always a small amount of non-specific binding (NSB) of label to the walls of the reaction tube. NSB is assessed by setting up a tube containing label but no sample or antibody which is then carried through the entire assay. The count rate for the NSB tube is subtracted from that of the TB tube in order to obtain a total binding value:

$$B_0 = TB - NSB \quad \text{..... Eq. 15.15}$$

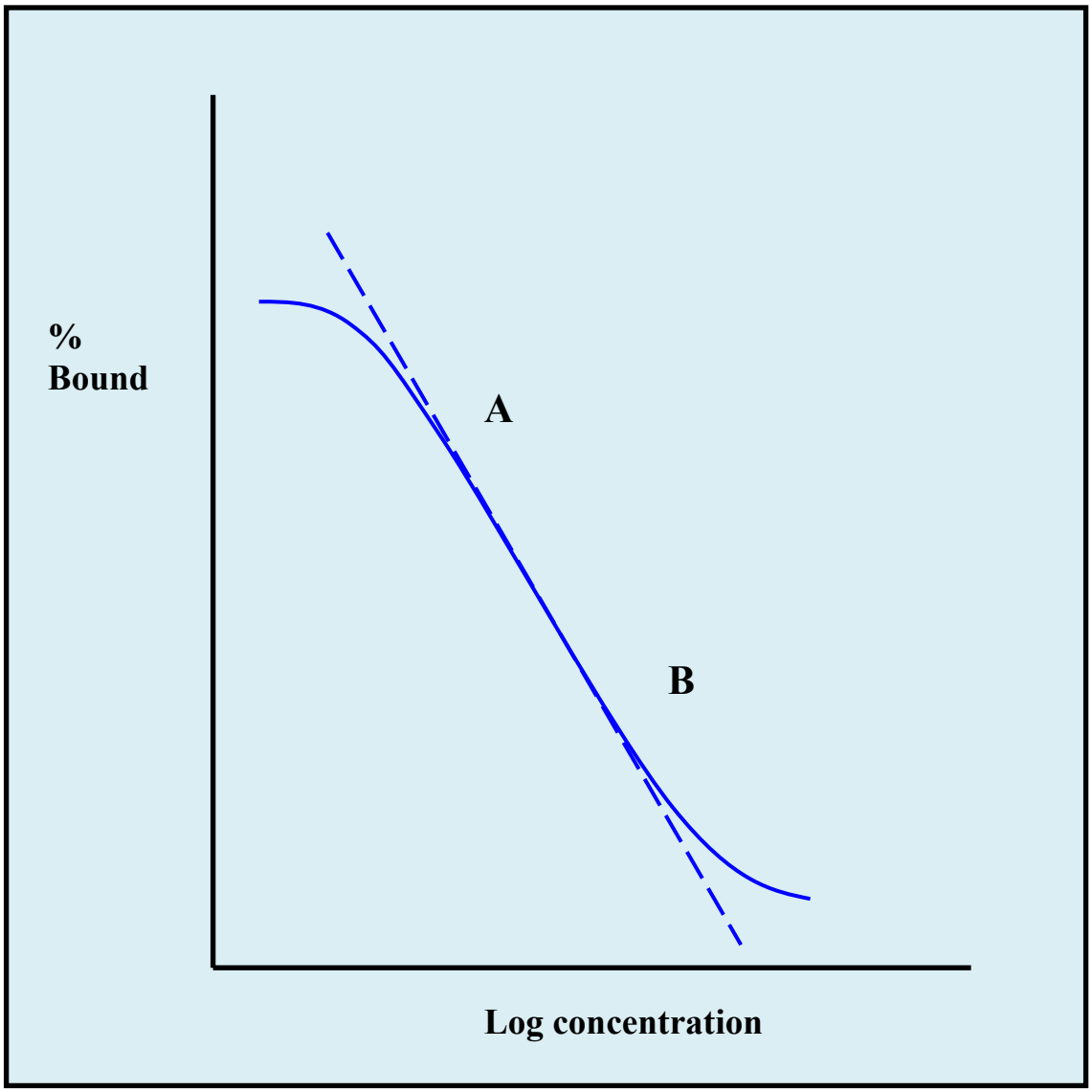


Figure 15.4 **Schematic diagram for the dose response curve of a typical radioimmunoassay. The area between A and B is the analytically useful range of the assay**

The *NSB* counts are also subtracted from the count rate for each standard and sample before expression as a percentage of B_0:

$$B/B_0\ (\%) = \frac{(\text{Standard/sample cpm} - NSB) \times 100}{B_0} \quad \ldots\ldots\ \textbf{Eq. 15.16}$$

A standard curve is then plotted of $B/B_0(\%)$ versus the logarithm of concentration. Values for unknowns are then obtained from the standard curve in the usual way (remembering to take the antilogarithm of the result). Nowadays, computer programmes are usually used to carry out the entire calculation including fitting the best curve to the standard data.

It is customary to set up a tube containing label alone which is not taken through the entire assay procedure but used to obtain a measure of the total count rate (*TC*). This, together with the B_0 result is a useful quality control tool to ensure that the antibody is able to bind a reasonable amount of label.

Question Q 15(10)

The following data were obtained for a plasma cortisol radioimmunoassay using second antibody separation. Calculate the cortisol concentration in the serum sample.

Sample	**Primary Antibody**	**Label**	**2nd antibody**	**Average cpm**
None	-	+	-	12,500
Buffer	-	+	+	200
Buffer	+	+	+	8,050
50 nmol/L standard	+	+	+	4,250
100 " "	+	+	+	3,720
200 " "	+	+	+	3,190
400 " "	+	+	+	2,675
800 " "	+	+	+	2,120
1600 " "	+	+	+	1,600
3200 " "	+	+	+	1,065
Serum	+	+	+	2,490

Answer Q 15(10)

First identify tubes to be used for total counts (*TC*), non-specific binding (*NSB*) and total binding (*TB*).

TC is the tube which contains label and nothing else (it is not used in the calculation of results). This is the first tube, therefore *TC* = 12,500 cpm.

NSB is the tube which contains label, buffer and second antibody and is taken through the entire assay procedure. This is the second tube so *NSB* = 200 cpm.

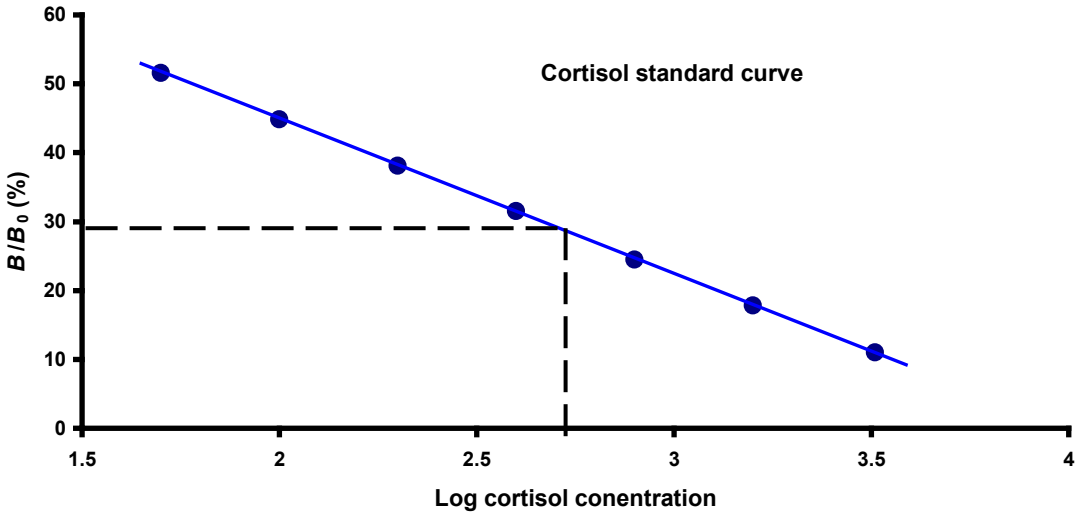
TB is the tube which contains everything except the analyte in question (i.e. buffer instead of sample). This is the third tube, therefore *TB* = 8,050 cpm.

Maximum binding (B_0) when analyte is absent is obtained by subtraction of *NSB* from *TB*:

$$B_0 = TB - NSB = 8{,}050 - 200 = 7{,}850 \text{ cpm.}$$

The *NSB* is also subtracted from each standard or sample count rate before it is expressed as a ratio to B_0 (See Eq. 15.16). This is best laid out in tabular form:

Sample	**Log_{10} conc**	**cpm**	**cpm – *NSB***	***B*/*B*$_0$(%)**
TC		12,500		
NSB		200		
TB		8,050	7,850	100
Standard 50 nmol/L	1.70	4,250	4,050	51.6
" 100 "	2.00	3,720	3,520	44.8
" 200 "	2.30	3,190	2,990	38.1
" 400 "	2.60	2,675	2,475	31.5
" 800 "	2.90	2,120	1,920	24.5
" 1600 "	3.20	1,600	1,400	17.8
" 3200 "	3.51	1,065	865	11.0
Serum		2,490	2,290	29.2



Note that the value for B_0 (100%) cannot be plotted since the log of zero (concentration for *TB*) has no meaning.

Serum sample B/B_0 (%) = 29.2.

From standard curve corresponding log concentration = 2.7

Therefore cortisol = antilog 2.7 = **500 nmol/L**

ADDITIONAL QUESTIONS

1. A 0.5 mL sample of urine is extracted into dichloromethane. An aliquot of the extract is analysed by HPLC and found to give an apparent original concentration of 320 nmol/L of analyte Y. 100 µL of Y standard with a concentration of 880 nmol/L is added to a further 0.5 mL sample of the same urine and the sample mixed. 0.5 mL of the mixed sample is then processed as before, giving a measured concentration of 405 nm/L. Calculate the recovery of analyte Y.

2. A new method for HCG in urine is being evaluated. The concentration in a sample from a pregnant woman is measured at 8240 IU/L. A 50 µL aliquot of an international standard containing 50,000 IU/L is added to 450 µL of the same urine sample and the sample mixed. On measuring the mixed sample, the new concentration is found to be 12100 IU/L. What is the recovery of HCG by this method?

3. Measurement of plasma AFP is used to monitor a patient with a teratoma. If the initial concentration was 10,200 U/L what plasma level would you expect to find 21 days after successful surgery? Assume the half-life of AFP is 5.5 days.

4. A radioisotope has a half-life of 21 days. How long will it take for the activity to fall to 10% of the initial value?

5. In normal pregnancy serum beta hCG has a doubling time of approximately 2 days. How long will it take for the serum level to increase ten-fold?

6. A patient receiving parenteral nutrition is receiving 11.8 g nitrogen/24 h as amino acids. Urinary urea excretion is 580 mmol/24 h. Indicating what assumptions you make, calculate whether she is in positive or negative nitrogen balance.

7. A 30 min basal gastric secretion sample (total volume 27 mL) required 2.5 mL of 0.1 M NaOH to titrate 5 mL of the material to pH 7.4. Calculate the basal acid secretion rate in mmol/h.

8. A five day faecal fat collection was homogenised and diluted to 1500 mL. A 10 mL aliquot of the homogenate was subjected to hydrolysis and the fatty acids were extracted. The volume of 0.05 M sodium hydroxide required to effect neutralisation was 48 mL. Calculate the fat excretion in mmol/24 h.

9. Gas chromatography for a drug involves adding equal amounts of internal standard to standard or sample prior to analysis. The following peak areas were obtained:

Sample	Peak area	
	Internal standard	Drug
Standard (200 nmol/L)	50,000	200,000
Patient	40,000	150,000

Calculate the drug concentration in the sample.

10. Genotyping of a group of 100 unrelated individuals for a two-allele polymorphism showed that the allele frequencies were:

A 0.65
B 0.35

Calculate the expected percentages of heterozygotes (AB) and homozygotes (AA and BB) in the group.

11. The prevalence of an inherited metabolic disease (inherited in an autosomal recessive manner due to a single allele) is 1 in 2,500. A survey identified 1 in 50 of the population as asymptomatic carriers. Is this finding consistent with a population in a Hardy-Weinberg equilibrium?

12. The following data were obtained for a digoxin radioimmunoassay employing PEG precipitation of the primary antibody. The assay was performed in duplicate. Calculate the digoxin concentration in the serum sample.

Sample	Duplicate cpm	
	1	**2**
TC	15,100	15,900
NSB	320	380
TB	11,350	11,650
0.2 nmol/L standard	10,320	10,980
0.4 " "	9,250	8,340
0.8 " "	6,782	6,630
1.2 " "	5,104	5,890
2.4 " "	3,700	3,430
4.8 " "	1,350	1,650
Patient serum	4,350	5,000

Appendix I

Answers to further questions

Chapter 1

1. a) 1.25 g/L
   b) 250 mmol/L
   c) 0.000236 µmol/L
   d) 1600 ng/mL
2. a) 6.7 mmol/L
   b) 2.0 mmol/L
   c) 7.5 mmol/L
   d) 58 µmol/L
3. a) 360 mg/100 mL
   b) 6.4 mEq/L
   c) 86 mg%
   d) 2.8 mg%
4. a) 1.5 mmol/L
   b) 12.5 µmol/L
   c) 2.5 g/L
   d) 3.25×10^{-3} µmol/L
5. a) 0.30 mmol
   b) 25 µmol/min/250 mL
6. 3.0×10^{-10} mol/L

Chapter 2

1. 7 g
2. 1453 mmol/L
3. 0.25 g
4. 100 mL
5. 101
6. 8.8 mL
7. Potassium = 0.022 mol/L
   Sodium = 0.57 mol/L
   Chloride = 0.59 mol/L
8. 950 mL
9. 5.2 mmol/L
10. Phosphate = 100 mmol/L
    40 mL needed

Chapter 3

1. 0.86
2. 36 to 45 nmol/L
3. 25:1
4. 6.80
5. Na_2CO_3 = 7.58 g
   $NaHCO_3$ = 2.39 g
6. 39.8 g sodium lactate
   0.22 mL lactic acid
7. 49 mmol
8. 4.65

Chapter 4

1. 207 mL
2. a) 0.022
   b) 0.125
   c) 0.301
   d) 0.602
   e) 1
   f) 2
3. a) 79 %
   b) 56 %
   c) 32 %
   d) 18 %
   e) 10 %
   f) 1 %
4. 7.35 $L.mol^{-1}.cm^{-1}$
5. 0.069
6. 153 nmol/g dry wt
7. 68 %
   16.7 $L.mol^{-1}.cm^{-1}$
8. NADH = 53.5 µmol/L
   NAD = 23.7 µmol/L
9. 97 %
10. Serum creat = 75 µmol/L
    Urine creat = 7.5 mmol/L
11. Linear to 15 mmol/L
    12.5 mmol/L

Chapter 5

1. 15.7 mmol/24 h
2. 14.3 mmol/12 h
3. Clearance = 170 mL/min
   ? >24h collection
4. Filtration = 1584 mg/24 h
   ? tubular reabsorption
5. 500 mL/min?
   ? tubular secretion
6. 36 mmol/L
7. Increased Na excretion
   by 192 mmol/24 h
8. 28 g Na (or 71 g NaCl)
9. 0.069 (= 6.9%)
10. 40 mL/min/1.73 m^2
11. 9 mmol/L glomerular filtrate
12. 0.51 mL/min
13. GFR = 39 mL/min/ 1.73 m^2
    Clearance = 28 mL/min
    ? incomplete urine collection
    Failure to correct clearance to body surface area
    Tubular secretion of creatinine

Chapter 6

1. 293 mOsm/kg
2. 857 mOsm/kg
3. 396 mOsm
4. 20 mOsm/kg
5. a) Gap = 52 mOsm/kg
   b) Ethanol = 59 mOsm/kg
   Ethanol explains gap.

Chapter 7

1. 40 h
2. a) 39 L
   b) 20 h
3. 1.25 mg/L (total body water)
   3.75 mg/L (ECF only)
4. a) 1.5 h
   b) 59 L
5. 2.0 h (for 100 nmol/L)
   2.9 h (for 75 nmol/L)
6. 6.6 h
7. 400 mg

Chapter 8

1. Positive by 500 mL
2. 6.6 L (or 5L)
3. Na^+ decrease by 3 mmol/L
4. 143 mmol/L

Chapter 9

1. 276 IU/L
2. a) 113 IU/L
   b) 1.88 x 10^{-6} kat/L
3. 1.85
4. 476
5. 0.8 V_{max}
6. a) K_m = 8.0 x 10^{-8} mol/L
   b) V_{max} = 1.9 x 10^{-7} mol/min
   c) V_{max} = 1.5 x 10^{-7} mol/min
   K_m = 1.5 x 10^{-5} mol/L
7. a) 1/[S] = 10^3 L/mol, 1/v = 10^6 min/mol
   b) [S]/v = 10^3 min/L, [S] = 10^{-3} mol/L
   c) v = 10^{-6} mol/min, v/[S] = 10^{-3} L/min
8. Competitive, K_i = 2.6 x 10^{-5} mol/L
9. K_m @ pH 7.4 = 6.2 x 10^{-3} mol/L
   K_m @ pH 5.5 = 1.8 x 10^{-3} mol/L
   Assuming equilibrium conditions
   highest affinity at pH 5.5
10. Uncompetitive, K_i approx 4 x 10^{-3} mol/L

Chapter 10

1. Mean = 70.25 g/L
   Variance = 7.82 g/L
   SD = 2.80 g/L
   CV = 4.0%
   95% limits = 64.8 - 75.7 g/L
2. 10
3. Least sig change = 14.8%
   Actual change = 14.5%
   Not significant
4. 0.74
5. Less than 6 mU/L
   (0.50-5.51 mU/L)
6. a) ± 0.36 mL
   b) ± 0.11 mL
7. 11.8%

Chapter 11

1. m = 101.1
   s = 10.6
   SE_m = 3.4
2. t = -1.75 (Not sig)
   F = 2.25 (Not sig)
3. 0.01
4. Paired t = -1.98
   Not significant
5. F = 2.34
6. Not significant

Chapter 12

1. a) 17.5 mmol/L
   b) 166 mmol/L
2. OLD =
   (NEW + 41)/0.65
3. r = 0.30 (Not sig)
4. ? relationship linear
   ? $s_{residual}$ (s_{yx})

Chapter 13

1. a) PV(+) = 0.95 or 95%
   PV(-) = 0.95 or 95%
   b) PV(+) = 0.004 or 4%
   PV(-) = 1.00 or 100%
2. a) 16.5 (2 sig figs)
   b) 1990
   c) Metanephrines since no false negatives
3. a) 0.49 or 49%
   b) 0.98 or 98%
4. SENS = 0.84 or 84%
   SPEC = 0.92 or 92%
   PV+ = 0.86 or 86%
5. Prev = 40 % PV(+) = 0.84 or 84%
   Prev = 0.4 % PV(+) = 0.030 or 3.0%
6. a) PV(+) = 0.33 or 33%
   b) Pre-test odds = 0.053
   c) LR+ = 9.4
   d) Post-test odds = 0.50
   e) Post-test prob = 0.33 or 33%
7. 0.48 or 48%

Chapter 14

1. 0.88 or 88% (2 sig figs)
2. 95

Chapter 15

1. 94%
2. 94%
3. 723 U/L
4. 70 days
5. 6.6 days
6. Negative (-4.4 g/24 h or -9.7 g/24 h)
7. 2.7 mmol/h
8. 24 mmol fat/24 h as triglyceride
9. 188 nmol/L
10. Heterozygotes 45.5%
    Homozygotes 54.5%
11. Consistent
    $X^2 = 0.010$ (not sig)
12. Digoxin = 1.6 nmol/L

Worked Answers to Further Questions

Chapter 1

(Atomic weights: C = 12; H = 1; O = 16; Ca = 40; N = 14)

1. *Convert the following: a) 125 mg% to g/L; b) 0.25 mol/L to mmol/L; c) 0.236 nmol/L to µmol/L; d) 1.6 mg/L to ng/mL.*

a) *Convert 125 mg% to g/L*

125 mg% is the same as 125 mg/100 mL

Multiply by 10 to convert volume from 100 mL to 1000 mL (i.e. 1 L)
Divide by 1000 to convert from mg to g (there are 1000 mg in one g)

$$125 \text{ mg\%} = \frac{125 \times 10}{1000} = \textbf{1.25 g/L}$$

b) Convert 0.25 mol/L to mmol/L

There are 1000 mmol in one mol. Therefore multiply by 1000.

$$0.25 \text{ mol/L} = 0.25 \times 1000 = \textbf{250 mmol/L}$$

c) Convert 0.236 nmol/l to µmol/L

One nmol = 1.0×10^{-9} mol, one µmol = 1.0×10^{-6} mol.
Therefore one µmol = 1.0×10^{3} nmol = 1000 nmol
Division of 1 nmol/L by 100 converts to µmol/L

$$0.236 \text{ nmol/L} = \frac{0.236}{1000} = \textbf{0.000236 µmol/L}$$

d) Convert 1.6 mg/L to ng/mL

There are 1,000,000 (or 1.0×10^6) ng in 1 mg

There are 1000 mL in one L

Therefore multiplication by 1,000,000 and division by 1000 converts from mg/L to ng/mL:

$$1.6 \text{ mg/L} = \frac{1.6 \times 1{,}000{,}000}{1{,}000} = \textbf{1600 ng/mL}$$

2. *Convert the following concentrations to "SI" units: a) plasma glucose 120 mg%; b) serum calcium 4.0 mEq/L; c) BUN 21 mg%; d) Serum creatinine 0.66 mg%.*

a) Convert plasma glucose = 120 mg% to SI units

'SI' units for glucose are mmol/L

120 mg% can also be written 120 mg/100mL

$$\text{Concentration (mmol/L)} = \frac{\text{Concentration (mg/L)}}{\text{Molecular weight}}$$

Multiplication of concentration in mg/100 mL by 10 converts to mg/L

Formula of glucose = $C_6H_{12}O_6$

Atomic wt carbon	= 12, therefore C_6	=	6 x 12	=	72
Atomic wt hydrogen	= 1, therefore H_{12}	=	12 x 1	=	12
Atomic wt oxygen	= 16, therefore O_6	=	6 x 16	=	96
	Molecular weight of glucose	=			180

$$\text{Therefore 120 mg\% glucose} = \frac{120 \times 10}{180} = \textbf{6.7 mmol/L} \text{ (2 sig figs)}$$

b) Convert serum calcium from 4.0 mEq/L to 'SI' units

The SI units for calcium are mmol/L

Concentration (mEq/L) = Concentration (mmol/L) x valency

Calcium ions are divalent (i.e. Ca^{++}) so that the valency is 2

Therefore, serum calcium (mmol/L) = $\frac{4.0}{2}$ = **2.0 mmol/L**

c) Convert BUN 21 mg% to urea 'SI' units

SI units for serum urea = mmol/L

mg% can also be written mg/100 mL

Multiplication of BUN mg% by 10 converts to mg/L (since 1 L = 1000 mL)

Division of blood urea nitrogen (BUN) in mg% by the molecular weight of nitrogen gives the blood urea nitrogen in mmol/L.

The formula of molecular nitrogen is N_2. The atomic weight of nitrogen is 14.

Molecular weight of nitrogen (N_2) = 2 x 14 = 28

The formula for urea is $CO(NH_2)_2$. Therefore each mol of urea contains one mol of nitrogen (N_2).

Therefore, serum urea (mmol/L) = $\frac{\text{BUN (mg\%) x 10}}{28}$

Serum urea (mmol/L) = $\frac{21 \text{ x } 10}{28}$ = **7.5 mmol/L**

d) Convert serum creatinine 0.66 mg% to 'SI' units

The SI units for creatinine are µmol/L

There are 1000 µg in one mg so that multiplication by 1000 converts from mg% to µg%

µg% can also be written µg/100 mL
Multiplication by 10 converts from µg/100 mL to µg/L (1 L = 1000 mL)

Division by the molecular weight of creatinine converts from µg/L to µmol/L. Formula of creatinine is: $C_4H_7ON_3$

Carbon atomic wt	= 12,	C_4	= 4 x 12	=	48
Hydrogen atomic wt	= 1,	H_7	= 7 x 1	=	7
Oxygen atomic weight	= 16,	O	= 1 x 16	=	16
Nitrogen atomic wt	= 14,	N_3	= 3 x 14	=	42
			Creatinine molecular weight	=	113

$$\text{Creatinine (µmol/L)} = \frac{\text{Creatinine (mg\%) x 10 x 1000}}{\text{Molecular weight}}$$

$$\text{Creatinine (mmol/L)} = \frac{0.66 \times 10 \times 1000}{113} = \textbf{58}\ \text{µmol/L (2 sig figs)}$$

3. *Convert the following: a) plasma glucose from 20 mmol/L to mg/100 mL; b) serum calcium from 3.2 mmol/L to mEq/L; c) serum urea from 30.6 mmol/L to mg% BUN; d) serum creatinine from 250 µmol/L to mg%.*

a) Convert plasma glucose 20 mmol/L to mg/100 mL

Division by 10 converts from mmol/L to mmol/100 mL (since I L = 1000 mL)

Multiplication by the molecular weight of glucose converts from mmol/100 mL to mg/100 mL. Formula of glucose is $C_6H_{12}O_6$.

Atomic wt carbon = 12, therefore C_6 = 6 x 12 = 72
Atomic wt hydrogen = 1, therefore H_{12} = 12 x 1 = 12
Atomic wt oxygen = 16, therefore O_6 = 6 x 16 = 96

Molecular weight of glucose = 180

$$\text{Therefore, glucose (mg/100mL)} = \frac{\text{Glucose (mmol/L) x 180}}{10}$$

$$\text{Glucose (mg/100 mL)} = \frac{20 \times 180}{10} = \mathbf{360\ mg/100mL}$$

b) Convert serum calcium from 3.2 mmol/L to mEq/L

Concentration (mEq/L) = Concentration (mmol/L) x valency

Calcium ions are divalent (i.e. Ca^{++}) so that the valency is 2

Therefore, calcium (mEq/L) = calcium (mmol/L) x 2

Calcium (mEq/L) = 3.2 x 2 = **6.4 mEq/L**

c) Convert serum urea from 30.6 mmol/L to mg% BUN

Division by 10 converts from mmol/L to mmol/100 mL (since 1 L = 1000 mL), which can also be written mmol%.

Since the formula of urea is $CO(NH_2)_2$, each mol contains 1 mol of molecular nitrogen (N_2). The atomic weight of nitrogen is 14, so its molecular weight (N_2) is 2 x 14 = 28. Therefore, multiplication of urea in mmol% by 28 gives the BUN in mg%.

$$\text{Therefore, BUN (mg\%)} = \frac{\text{Urea (mmol/L) x 28}}{10}$$

$$\text{BUN (mg\%)} = \frac{30.6 \times 28}{10} = \mathbf{86\ mg\%}\ \text{(2 sig figs)}$$

d) Convert serum creatinine from 250 µmol/L to mg%

Division by 10 converts from µmol/L to µmol/100 mL (since 1 L = 1000 mL), which can also be written as µmol%.

Division by 1000 converts from µmol% to mmol% (since 1000 µmol = 1 mmol).

Multiplication by the molecular weight of creatinine converts from mmol% to mg%. Creatinine has the formula $C_4H_7ON_3$.

Carbon atomic wt	=	12,	C_4	=	4 x 12	= 48
Hydrogen atomic wt	=	1,	H_7	=	7 x 1	= 7
Oxygen atomic weight	=	16,	O	=	1 x 16	= 16
Nitrogen atomic wt	=	14,	N_3	=	3 x 14	= 42
					Creatinine molecular weight	= 113

$$\text{Therefore, creatinine (mg\%)} = \frac{\text{creatinine (µmol/L) x 113}}{10 \times 1000}$$

$$\text{Creatinine (mg\%)} = \frac{250 \times 113}{10 \times 1000} = \textbf{2.8 mg\%} \text{ (2 sig figs)}$$

4. *Convert the following: a) 1.5 x 10^{-3} M to mmol/L; b) 1.25 x 10^{-5} M to µmol/L; c) 2.5 x 10^{2} mg/100 mL to g/L; d) 3.25 x 10^{-6} mmol/L to µmol/L.*

a) Convert 1.5 x 10^{-3}M to mmol/L

'M' is an abbreviation for mol/L. There are 1000 mmol in one mol.

Therefore, multiplication of a concentration in mol/L by 1000 converts it to mmol/L.

1.5 x 10^{-3} is the same as $\frac{1.05}{10^3}$

10^3 means 10 multiplied by itself 3 times i.e. 10 x 10 x 10 = 1000

Therefore, $1.5 \times 10^{-3} = \frac{1.5}{1000} = 0.0015$

Another way of looking at it is that the power minus 3 means that we must move the decimal point 3 places to the left (as opposed to a positive power which would have meant moving it 3 places to the right). Moving the decimal point one pace to the left gives 0.15, 2 places gives 0.015 and 3 places gives 0.0015.

Combining these moves:

1.5×10^{-3} M = 0.0015 x 1000 = **1.5 mmol/L**

In other words whenever we see a molar concentration with the term ' x 10^{-3}, it is really the same as the concentration in mmol/L.

Another way to approach this problem is that multiplying by 1000 to convert from mol to mmol is the same as multiplying by 10^3 in which case the calculation becomes:

$1.0 \times 10^{-3} \times 10^3 = 1.0$

since the 10^{-3} and 10^3 cancel (i.e. we move the decimal point 3 places to the left then back 3 places to the right to the original position).

b) *Convert 1.25 x 10^{-5} M to µmol/L*

The symbol 'M' stands for mol/L.

Multiplication by 1,000,000 converts from mol to µmol (since there are 1,000,000 µmol in one mol). 1,000,000 can also be written as 10^6

Therefore 1.25×10^{-5} M = $1.25 \times 10^{-5} \times 10^6$ µmol/L

Since $10^{-5} \times 10^6 = 10^1$ (which is simply 10) the calculation becomes:

1.25×10^{-5} M = 1.25 x 10 = **12.5 µmol/L**

This is the same as moving the decimal point 5 places to the left, then 6 places

to the right i.e. a net movement of one place to the right.

c) *Convert 2.5 x 10^2 mg/100 mL to g/L*

Multiplication by 10 converts from mg/100 mL to mg/L (since 1 L = 1000 mL).

Division by 1000 converts from mg/L to g/L (since there are 1000 mg in one g)

2.5 x 10^2 mg/100 mL means 2.5 multiplied by 10 squared (i.e. 100) = 250 mg/100 mL

Therefore, 2.5 x 10^2 mg/100 mL = $\frac{250 \times 10}{1000}$ = **2.5 g/L**

Another way of doing this is that we first move the decimal point 2 places to the right (the same as multiplying by 10^2), a further one place to the right to convert from 100 mL to 100 mL (making 3 moves to the right altogether). A further 3 moves to the left (the same as dividing by 1000 to convert from mg to g) takes us back to the starting position and an answer of 2.5 g/L.

d) *Convert 3.25 x 10^{-6} mmol/L to µmol/L*

Multiplication of mmol/L by 1000 converts to µmol/L (since 1 mmol = 1000 µmol). Multiplication by 1000 is the same as multiplication by 10^3

Therefore, 3.25 x 10^{-6} mmol/L = 3.25 x 10^{-6} x 10^3 = **3.25 x 10^{-3} µmol/L**

3.25 x 10^{-3} µmol/L can also be written 0.00325 µmol/L.

(The same result is obtained by moving the decimal point 6 places to the left then back 3 places to the right).

5. *After incubation of an enzyme with substrate for 30 min the concentration of product in the reaction mixture was 3.00 x 10^{-3} M. a) How many mmol of product would be present in 100 mL of the reaction mixture; and b) what is the rate of formation of product in 250 mL of reaction mixture expressed as µmol/min?*

   a) *How many mmol of product is present in 100 mL of reaction mixture?*

   If the concentration in the reaction mixture is 3.00 x 10^{-3} M then each litre contains 3.0 x 10^{-3} mol of product.

   Division by 10 gives the number of mol in 100 mL (since 1 L = 1000 mL)

   Multiplication by 1000 converts from mol to mmol (since 1 mol = 1000 mmol)

   3.00 x 10^{-3} can also be written as $\frac{3.00}{10^3}$ or $\frac{3.00}{1000}$

   Therefore concentration (mmol/100mL) = $\frac{\text{concentration (mol/L) x 1000}}{10}$

   Amount of product in 100 mL = $\frac{3.00 \text{ x } 1000}{10 \text{ x } 1000}$ = **0.30 mmol**

   b) *What is the rate of formation of product in 250 mL of reaction mixture expressed as µmol/min?*

   3.00 x 10^{-3} M means 3.00 x 10^{-3} mol/L

   Multiplication by 1,000,000 converts from mol/L to µmol/L (since 1 mol = 1,000,000 µmol). 1,000,000 can also be written 10^6.

   Division by 4 gives the amount of product present in 250 mL (since 1000/4 = 250).

   Since this amount of product was formed over 30 min, division by 30 gives the amount which would be formed in one minute.

Therefore number of µmol formed in 250 mL in one minute

$$= \frac{\text{Molar concentration} \times 10^6}{4 \times 30}$$

$$= \frac{3.00 \times 10^{-3} \times 10^6}{4 \times 30} = \mathbf{25\ \mu mol/min/250\ mL}$$

NB $10^{-3} \times 10^6 = 10^3$ i.e. move decimal point 3 places to the left, then 6 places to the right making 3 places to the right overall, which is the same as multiplying by 1000.

6. *An acid dissociates in solution to give its conjugate base and hydrogen ions. Calculate the dissociation constant if urine contains 0.1 M of undissociated acid, 25 x 10^{-5} mol/L of its conjugate base and 120 nmol/L of hydrogen ions? NB the dissociation constant is the product of the concentrations of conjugate base and hydrogen ions divided by the concentration of undissociated acid.*

The dissociation of the acid can be written:

Acid	↔	Conjugate base$^-$	+	H^+
0.1M		25 x 10^{-5} mol/L		120 nmol/L

Before calculating the dissociation constant (K) convert all concentrations to the same units. It doesn't really matter which units but conventionally molar concentrations (mol/L) are used.

0.1 M undissociated acid is the same as 0.1 mol/L of acid, which can also be written as 1.0 x 10^{-1} mol/L.

There are 1,000,000,000 (i.e. 10^9) nmol in one mol so that 120 nmol/L hydrogen ions can be written 120 x 10^{-9} mol/L or more conveniently 1.20 x 10-^{7} mol/L (moving the decimal point 2 places to the left and decreasing the power of 10 by 2 from -9 to -7).

The conjugate base concentration is already in mol/L, but would be more conventionally written as 2.5 x 10^{-4} (reduction from 25 to 2.5 involves moving the decimal point one place to the left so that the power of 10 must be increased by one i.e. changed from -5 to -4).

The expression for the dissociation constant, with molar concentrations in square brackets [] can be written:

$$K = \frac{[\text{Conjugate base}]\ (\text{mol/L}) \times [\text{Hydrogen ions}]\ (\text{mol/L})}{[\text{Acid}]\ (\text{mol/L})}$$

An "expression" for the units of K can be written:

$$\frac{(\text{mol/L}) \times (\text{mol/l})}{(\text{mol/L})}$$

One set of (mol/) above the line cancels with (mol/L) below the line leaving **mol/L** as the units.

Calculation of K from these data gives:

$$K = \frac{2.5 \times 10^{-4} \times 1.20 \times 10^{-7}}{1.0 \times 10^{-1}} = 2.5 \times 1.20 \times 10^{-10} = \mathbf{3.0 \times 10^{-10}\ mol/L}$$

Chapter 2

(Atomic weights: H = 1; C = 12; O = 16; P = 31; Na = 23; K = 39; Ca = 40; S = 32; Cl = 35.5)

1. *How many grams of albumin are required to prepare 100 mL of a solution containing 70 g/L.*

   70 g/L is the same as 70 g/1000 mL (since 1L = 1,000 mL)

   If 1000 mL contain 70 g of albumin then each mL contains $\frac{70 \text{ g}}{1{,}000}$

   And 100 mL must conation 100 times this amount, so that the weight required to make 100mL of 70 g/L albumin

   $$= \frac{70 \times 100}{1{,}000} = \mathbf{7\ g}$$

   Another way of looking at this is that 100 mL is one tenth of 1 L (i.e. 1,000 mL) so that one tenths of the amount present in l L is required.

2. *Calculate the concentration of sodium ions (in mmol/L) in a solution prepared by dissolving 85 g of sodium chloride in 1 litre of water.*

   Concentration of sodium chloride = 85 g/L

   Division by the molecular weight of sodium chloride gives the concentration in mol/L:

   $$\text{Concentration (mol/L)} = \frac{\text{Concentration (g/L)}}{\text{Molecular weight}}$$

   Multiplication by 1,000 converts from mol/L to mmol/L (since 1 mol = 1,000 mmol)

Formula for sodium chloride = NaCl

Atomic weight of Na = 23; atomic weight of Cl = 35.5

Therefore molecular weight of NaCl = 23 + 35.5 = 58.5

Therefore, NaCl (mmol/L) = $\frac{\text{NaCl (g/L)} \times 1{,}000}{58.5}$

Since each molecule of NaCl dissociates to give one ion of Na^+, this is also the concentration of sodium ions in mmol/L.

Concentration of Na^+ (mmol/L) = $\frac{85 \times 1{,}000}{58.5}$ = **1453 mmol/L**

3. *What weight of calcium carbonate must be dissolved in 500 mL of dilute acid to provide a calcium standard containing 5.0 mmol/L ?*

Calcium carbonate has the formula $CaCO_3$ so that each mol contains 1 mol of calcium. Therefore, the standard solution will need to contain 5.0 mmol/L of $CaCO_3$.

Atomic wt calcium = 40, therefore Ca = 1 x 40 = 40
Atomic wt carbon = 12, therefore C = 1 x 12 = 12
Atomic wt of oxygen = 16, therefore O_3 = 3 x 16 = 48

Molecular weight of $CaCO_3$ = 100

1 L of 1 mol/L will contain 100 g of $CaCO_3$

1 L of 1mmol/L will contain $\frac{100}{1{,}000}$ g of $CaCO_3$ (since 1 mol/L = 1,000 mmol/L)

1 L of 5.0 mmol/L will contain $\frac{100 \times 5.0}{1{,}000}$ g $CaCO_3$

500 mL of 5.0 mmol/L will contain $\frac{100 \times 5.0}{1{,}000 \times 2}$ g $CaCO_3$ (1L = 1,000mL)

= **0.25 g** (= 250 mg)

4. *A solution contains 5% sucrose. How much of this solution would you dilute to prepare 500 mL of % sucrose?*

The total amount of sucrose (as opposed to concentration) remains the same after dilution. The amount of sucrose in a given volume of solution is equal to the volume multiplied by concentration. Therefore the following expression can be written:

Initial volume x Initial concentration = Final volume x Final concentration

The units for volume and concentration must be the same on both sides of the equation.

Initial volume = unknown
Initial concentration = 5%
Final volume = 500 mL
Final concentration = 1%

Substituting these values the initial volume can be calculated:

Initial volume (mL) x 5 = 500 x 1

$$\text{Initial volume (mL)} = \frac{500 \times 1}{5} = \textbf{100 mL}$$

Another way to do this is that the final concentration (1%) is one fifth of the initial value (5%) so that the 5% sucrose solution has to be diluted 5-fold. The volume required is 500 mL so that one fifth of this, 100 mL has to be diluted.

5. *50 µL of urine is added to 5 mL of water. What is the resulting dilution of the urine?*

Both volumes must be expressed in the same units. Multiplication of the volume of water (5 mL) by 1000 gives its volume in µL (5,000 µL) since 1 mL = 1,000 µL.

The total volume of diluted urine is the sum of the volumes of urine and water:

Final volume (µL) = 50 + 5,000 = 5,050 µL

The dilution is the number of times the urine was diluted which is the final volume divided by the initial volume:

$$\text{Dilution} = \frac{\text{Final volume}}{\text{Initial volume}} = \frac{5{,}050}{50} = \mathbf{101}$$

N.B. Concentration is the reciprocal of dilution. In this case 1/101 = 0.0099.

6. *Concentrated sulphuric acid (SG 1.84) is 96% by weight H_2SO_4. Calculate the volume of concentrated acid required to prepare 1 L of 0.1M H_2SO_4.*

First calculate the molecular weight of sulphuric acid:

Atomic wt of hydrogen	=	1,	2H	=	2 x 1	= 2
Atomic wt of sulphur	=	32,	S	=	1 x 32	= 32
Atomic wt of oxygen	=	16,	4O	=	4 x 16	= 64
					Molecular weight H_2SO_4	= 98

Therefore 1 L of 1 mol/L requires 98 g H_2SO_4

1 L of 0.1 mol/L requires $\frac{98}{10}$ g H_2SO_4 (since it is $1/10^{th}$ the strength of 1 mol/L)

Since the sulphuric acid is only 96% pure this weight must be multiplied by 100/96:

$$\text{Weight } H_2SO_4 \text{ required} = \frac{98 \times 100}{10 \times 96} = 10.21 \text{ g (4 sig figs)}$$

The volume required can be calculated from the specific gravity:

$$\text{Specific gravity (SG)} = \frac{\text{Weight}}{\text{Volume}}$$

$$\text{Volume} = \frac{\text{Weight}}{\text{Specific gravity}}$$

We are told that the specific gravity of H_2SO_4 is 1.16 so that 1 mL weighs 1.16 g. The volume which weighs 102.08 g is calculated as follows:

$$\text{Volume (mL)} = \frac{10.21}{1.16} = \mathbf{8.8\ mL}$$

7. *The following solutions were mixed together:*

50 mL potassium chloride (5.0 g/L)
100 mL sodium chloride (50 g/L)

Calculate the molar concentrations of potassium, sodium, and chloride ions.

First calculate the molar concentration of each individual solution:

Potassium chloride (KCl). Atomic wt K = 39, atomic wt Cl = 35.5
Molecular weight of KCl = 39 + 35.5 = 74.5

$$\text{KCl (mol/L)} = \frac{\text{KCl (g/L)}}{\text{Molecular wt}} = \frac{5.0}{74.5} = 0.0671\ \text{mol/L (3 sig figs)}$$

Sodium chloride (NaCl). Atomic wt Na = 23, atomic wt Cl = 35.5.
Molecular weight NaCl = 23 + 35.5 = 58.5

$$\text{NaCl (mol/L)} = \frac{\text{NaCl (g/L)}}{\text{Molecular wt}} = \frac{50}{58.5} = 0.855\ \text{(3 sig figs)}$$

For potassium:

50 mL 0.0671M KCl + 100 mL 0.855 M NaCL → 150 mL mixture

0.0671 M KCl contains 0.06371 M K^+. 50 mL is diluted to 150 mL.

Final K^+ (mol/L) x Final vol (mL) = Initial K^+ (mol/L) x Initial vol (mL)

Final K^+ (mol/L) x 150 = 0.0671 x 50

$$\text{Final } K^+ \text{ (mol/L)} = \frac{0.0671 \times 50}{} = \mathbf{0.022\ mol/L}\ \text{(2 sig figs)}$$

150

For sodium:

100 mL NaCl 0.855 mol/L + 50 mL KCl → 150 mL mixture

Final Na^+ (mol/L) x Final vol (mL) = Initial Na^+ (mol/L) x Initial vol (mL)

Final Na^+ (mol/L) x 150 = 0.855 x 100

$$\text{Final } Na^+ \text{ (mol/L)} = \frac{0.855 \times 100}{150} = \mathbf{0.57\ mol/L}$$

For chloride:

100 mL NaCL 0.855 mol/L + 50 mL KCl 0.0671 → 150 mL mixture

Final Cl^- (mol/L) x Final vol (mL) = [Initial Cl^- from KCl (mol/L) x

Final Cl^- (mol/L) x 150 = [0.0671 x

vol KCl (mL)] + [Initial Cl^- from NaCl (mol/L) x vol NaCl (L)]

50] + [0.855 x 100]

$$\text{Final } Cl^- \text{ (mol/L)} = \frac{[(0.0671 \times 50) + (0.855 \times 100)]}{150}$$

$$= \frac{3.355 + 85.5}{150}$$

$$= \frac{88.855}{150}$$

= **0.59 mol/L** (2 sig figs)

8. *If you have available 650 mL of 9 % ethanol, what is the maximum volume of 65 % ethanol you could prepare?*

Initial vol (mL) x Initial concn (%) = Final vol (mL) x Final concn (%)

650 x 95 = Final vol (mL) x 65

$$\text{Final vol (mL)} = \frac{650 \times 95}{65} = \textbf{950 mL}$$

N.B. The expected volume of water to be added to the 95% ethanol (950 – 650 = 300 mL) will be insufficient because mixing an alcohol with water always results in some contraction of the total volume. Therefore further water should be added until a final volume of 950 mL is reached.

9. *In order to prepare 1 L of a stock standard solution containing 0.2 mol/L, the appropriate amount of sodium dihydrogen orthophosphate dihydrate should be weighed out. Due to an error, the same weight of anhydrous sodium dihydrogen orthophosphate used. Working standard was prepared by taking 5 mL of this stock standard and diluting it to 250 mL. What is the phosphate concentration (in mmol/L) of the working standard?*

First calculate the weight of sodium dihdrogen orthophosphate dihydrate ($NaH_2PO_4.2H_2O$) which should have been used. Adding individual atoms together gives the empirical formula: NaH_6PO_6

Atomic wt sodium	=	23,	therefore Na	=	1 x 23	= 23
Atomic wt hydrogen	=	1,	therefore 6H	=	6 x 1	= 6
Atomic wt phosphorus	=	31,	therefore P	=	1 x 31	= 31
Atomic wt oxygen	=	16,	therefore 6O	=	6 x 16	= 96

Molecular weight $NaH_2PO_4.2H_2O$ = 156

Therefore, 1L 1.0 mol/L contains 156 g $NaH_2PO_4.2H_2O$

so that, 1 L 0.2 mol/L contains $\frac{156}{5}$ g = 31.2 g $NaH_2PO_4.2H_2O$

Next calculate the molar concentration if this weight (31.2 g) of anhydrous sodium dihydrogen orthophosphate (NaH_2PO_4) was dissolved in 1 L of water.

Atomic wt sodium	=	23, therefore Na	=	1 x 23	=	23
Atomic wt hydrogen	=	1, therefore 2H	=	2 x 1	=	2
Atomic wt phosphorus	=	31, therefore P	=	1 x 31	=	31
Atomic wt oxygen	=	16, therefore 4O	=	4 x 16	=	64
				Molecular weight NaH_2PO_4	=	120

$$\text{Concentration (mol/L)} = \frac{\text{Concentration (g/L)}}{\text{Molecular weight}}$$

$$\text{Concentration (mol/L)} = \frac{31.2}{120} = 0.26 \text{ mol/L}$$

As a short cut the target concentration (0.2 mol/L) could be simply multiplied by the ratio of the molecular weight of $NaH_2PO_4.2H_2O$ to that of NaH_2PO_4:

Actual concentration (mol/L) =

$$\frac{\text{Target concentration (0.2 mol?L) x MW } NaH_2PO_4.2H_2O}{\text{MW } NaH_2PO_4}$$

$$= \frac{0.2 \times 156}{120} = 0.26$$

Working standard was prepared by diluting 5 mL of this stock standard to 250 mL.

Initial concn (mol/L) x Initial vol (mL) = Final concn (mol/L) x Final vol (mL)
0.26 x 5 = Final concn (mol/L) x 250

$$\text{Final concn (mol/L)} = \frac{0.26 \times 5}{250} = 0.0052 \text{ mol/L}$$

Multiplication by 1,000 (since there are 1,000 mmol in a mol) converts this concentration to mmol/L:

Working phosphate standard concentration = 0.0052 x 1,000 = **5.2 mmol/L**

10. *Solution A contains 12.0 g of anhydrous sodium dihydrogen phosphate per litre. What is the phosphate concentration expressed as mmol/L? What volume of solution A needs to be diluted to 1 L to give a phosphate concentration of 4 mmol/L.*

First calculate the molecular weight of anhydrous sodium dihydrogen phosphate (NaH_2PO_4):

Atomic wt sodium	=	23,	therefore	Na	=	1 x 23	= 23
Atomic wt hydrogen	=	1,	therefore	2H	=	2 x 1	= 2
Atomic wt phosphorus	=	31,	therefore	P	=	1 x 31	= 31
Atomic wt oxygen	=	16,	therefore	4O	=	4 x 16	= 64
						Molecular weight NaH_2PO_4	= 120

Next calculate the molar concentration of a solution containing 12 g/L:

$$\text{Concentration (mol/L)} = \frac{\text{Concentration (g/L)}}{\text{Molecular weight}}$$

$$\text{Concentration (mol/L)} = \frac{12}{120} = 0.1 \text{ mol/L}$$

Multiplication by 1,000 gives the concentration in mmol/L (since 1 mol = 1,000 mmol):

Phosphate concentration = 0.1 x 1,000 = **100 mmol/L**

To prepare 1 L of 4 mmol/L phosphate:

Initial vol (mL) x Initial conc (mmol/L) = Final vol (mL) x Final conc (mmol/L)

Initial vol (mL) x 100 = 1,000 x 4

$$\text{Initial vol (mL)} = \frac{1{,}000 \times 4}{100} = \textbf{40 mL}$$

Chapter 3

1. *What is the pH of 0.5 per cent (w/v) hydrochloric acid (assume complete dissociation, atomic weight Cl = 35.5)?*

First calculate molar concentration of 0.5% HCl:

0.5 %w/v = 0.5 g/100 mL = 0.5 x 10 = 5.0 g/L

MW HCl = 1 + 35.5 = 36.5

$$\text{Molar conc} = \frac{\text{g/L}}{\text{MW}} = \frac{5.0}{36.5} = 0.137 \text{ mol/L} \quad (3 \text{ sig figs})$$

Next calculate pH assuming complete dissociation of HCl:

$$pH = -\log_{10}[H^+]$$

Substitute $[H^+] = 0.137$ mol/L

$$pH = -\log_{10} 0.137 = -(-0.86) = \mathbf{0.86} \quad (2 \text{ sig figs})$$

2. *The reference range for blood pH is often quoted as 7.35 – 7.45. Express this range in terms of nannomoles of hydrogen ion per litre.*

$$pH = -\log_{10}[H^+]$$

Rearranging: $\log_{10}[H^+] = -pH$

Therefore $[H^+] = \text{antilog}_{10}(-pH)$

Substitute pH = 7.35:

$$[H^+] = \text{antilog}_{10}(-7.35) = 4.47 \times 10^{-8} \text{ mol/L}$$

$$= 44.7 \times 10^{-9} \text{ mol/L} = \mathbf{45\ nmol/L} \quad (2 \text{ sig figs})$$

Substitute pH = 7.45

$[H^+]$ = $\text{antilog}_{10}(-7.45)$ = 3.55×10^{-8} mol/L

= 35.5×10^{-9} mol/L = **36 nmol/L** (2 sig figs)

3. *If the pH of urine is 6.0 and of blood 7.40, what is the gradient of hydrogen ion concentrations across the tubular cell walls?*

pH = $-\log_{10}[H^+]$ therefore $[H^+] = \text{antilog}_{10}(-pH)$

For urine substitute pH = 6.0:

$[H^+]$ = $\text{antilog}_{10}(-6.0)$ = 1.0×10^{-6} mol/L

= $1{,}000 \times 10^{-9}$ mol/L = 1000 nmol/L

For blood substitute pH = 7.40:

$[H^+]$ = $\text{antilog}_{10}(-7.40)$ = 3.98×10^{-8}

= 39.8×10^{-9} mol/L = 40 nmol/L

Gradient = $\frac{[H^+] \text{ in urine}}{[H^+] \text{ in blood}}$ = $\frac{1000}{40}$ = **25:1**

Another way of approaching this problem is to use the fact that the ration of two values is equal to the antilog of the difference between their logarithms.

In other words, substitute $\text{antolog}_{10}(-pH)$ for [H+]:

Gradient = $\frac{[H^+]_{urine}}{[H^+]_{blood}}$ = $\text{antilog}_{10}\frac{(-pH_{urine})}{(-pH_{blood})}$

= $\text{antilog}_{10}\{-pH_{urine} - (-pH_{blood})\}$

= $\text{antilog}_{10}(pH_{blood} - pH_{urine})$

$= \text{antilog}_{10}\ (7.4 - 6.0)$

$= \text{antiolg}_{10}\ 1.4$

$= \mathbf{25}$ (2 sig figs)

4. *Determine the secondary dissociation constant of phosphoric acid if blood of pH 7.00 contains 12.85 mg disodium hydrogen orthophosphate and 6.88 mg sodium dihydrogen orthophosphate per 100 mL of plasma.*

The dissociation to be considered is:

$$H_2PO_4^- \leftrightarrow HPO_4^{2-} + H^+$$

$$pH = pKa + \log_{10} \frac{[Na_2HPO_4]}{[NaH_2PO_4]}$$

Rearranging: $$pKa = pH - \log_{10} \frac{[Na_2HPO_4]}{[NaH_2PO_4]}$$

Next calculate the molar concentration of each phosphate.

$$[Na_2HPO_4] = \frac{12.85 \times 10}{1{,}000 \times MW}$$

MW $Na_2HPO_4 = (2 \times 23) + 1 + 31 + (4 \times 16) = 142$

$$[Na_2HPO_4] = \frac{12.85 \times 10}{1000 \times 142} = 0.000905 \text{ mol/L}$$

$$[NaH_2PO_4] = \frac{6.88 \times 10}{1{,}000 \times MW}$$

MW $NaH_2PO_4 = 23 + (2 \times 1) + 31 + (4 \times 16) = 120$

$$[NaH_2PO_4] = \frac{6.88 \times 10}{1000 \times 120} = 0.000573 \text{ mol/L}$$

Next substitute these molar concentrations into the rearranged Henderson-Hasselbalch equation and solve for pKa:

$$pKa = 7.0 - \log_{10} \frac{0.000905}{0.000573}$$

$$pKa = 7.0 - \log_{10} 1.58 = 7.0 - 0.20 = \mathbf{6.80}$$

5. *What weight of anhydrous sodium carbonate and sodium bicarbonate would be required to prepare 500 mL of 0.2 M buffer pH 10.7 (pK_a HCO_3^- = 10.3)?*

The relevant dissociation is:

$$HCO_3^- \rightarrow H^+ + CO_3^{2-}$$

$$pH = pKa + \log_{10} \frac{[CO_3^{2-}]}{[HCO_3^-]}$$

Rearrange, substitute pH = 10.7 and pKa = 10.3, then calculate ratio:

$$\log_{10} \frac{[CO_3^{2-}]}{[HCO_3^-]} = pH - pKa = 10.7 - 10.3 = 0.4$$

$$\frac{[CO_3^{2-}]}{[HCO_3^-]} = \text{antilog } 0.4 = 2.51 \quad \ldots\ldots\ldots \text{(i)}$$

Since the total concentration of both bicarbonate and carbonate in the buffer is 0.2 mol/L:

$$0.2 = [HCO_3^-] + [CO_3^{2-}] \quad \ldots\ldots\ldots\ldots \text{(ii)}$$

$$[HCO_3^-] = 0.2 - [CO_3^{2-}]$$

Substitute for $[HCO_3^-]$ in (i) and solve for $[CO_3^{2-}]$:

$$\frac{[CO_3^{2-}]}{0.2 - [CO_3^{2-}]} = 2.51$$

$$[CO_3^{2-}] = (2.51 \times 0.2) - 2.51\,[CO_3^{2-}]$$

$$[CO_3^{2-}] + 2.51\,[CO_3^{2-}] = 0.502$$

$$3.51\,[CO_3^{2-}] = 0.502$$

$$[CO_3^{2-}] = \frac{0.502}{3.51} = 0.143 \text{ mol/L}$$

Substitute $[CO_3^{2-}] = 0.143$ into (ii) and solve for $[HCO_3^-]$:

$$0.2 = [HCO_3^-] + 0.143$$

$$[HCO_3^-] = 0.2 - 0.143 = 0.057 \text{ mol/L}$$

Calculate weights of both sodium carbonate and bicarbonate needed to prepare 500 mL of buffer:

$$\text{Weight required (g)} = \frac{\text{Molar concentration (mol/L)} \times \text{MW}}{2}$$

MW Na_2CO_3 = (2 x 23) + 12 + (3 x 16) = 106

MW $NaHCO_3$ = 23 + 1 + 12 + (3 x 16) = 84

Wt Na_2CO_3 = $\frac{0.143 \times 106}{2}$ = **7.58 g** (3 sig figs)

WT $NaHCO_3$ = $\frac{0.057 \times 84}{2}$ = **2.39 g** (3 sig figs)

6. *Isotonic sodium lactate, pH 7.4, is commonly administered intravenously to combat metabolic acidosis. How many ml of concentrated lactic acid (85% w/w, density 1.2) and how many grams of anhydrous sodium lactate would be used to prepare 2.5L of this solution (pK_a lactic acid = 3.86)?*

The relevant dissociation is:

$$LactH \rightarrow Lact^- + H^+$$

$$pH = pKa + \log_{10} \frac{[Lact^-]}{[LactH]}$$

Substitute pH = 7.4 and pKa = 3.86 then calculate ratio:

$$7.4 = 3.86 + \log_{10} \frac{[Lact^-]}{[LactH]}$$

$$\log_{10} \frac{[Lact^-]}{[LactH]} = 7.4 - 3.86 = 3.54$$

$$\frac{[Lact^-]}{[LactH]} = \text{antilog}_{10}\, 3.54 = 3467 \ldots\ldots..(i)$$

The concentration is not given but we are told that the solution must be isotonic. Assuming physiological osmolarity is 285 mmol/L:

$$285 = [LactH] + [Lact^-] + [Na^+]$$

Where the concentrations of LactH, $Lact^-$ and Na^+ are mmol/L.
Since the concentrations of $Lact^-$ and Na^+ are equal:

$$285 = [LactH] + 2\,[Lact^-] \ldots\ldots\ldots\ldots (ii)$$

Rearranging:

$$[LactH] = 285 - 2\,[Lact^-]$$

Substitute for [LactH] in (i) and solve for $[Lact^-]$:

$$\frac{[Lact^-]}{285 - 2\,[Lact^-]} = 3467$$

$$[Lact^-] = 3467\,(285 - 2\,[Lact^-])$$

$$[Lact^-] = 988095 - 6934\,[Lact^-]$$

$$[Lact\text{-}] + 6935\,[Lact\text{-}] = 988095$$

$$6935\,[Lact^-] = 988095$$

$$[Lact^-] = \frac{988095}{6935} = 142 \text{ mmol/L} \quad (3 \text{ sig figs})$$

Substitute $[Lact^-] = 142$ into (ii) and solve for [LactH]:

$$285 = [LactH] + (2 \times 142)$$

$$[LactH] = 285 - 284 = 1 \text{ mol/L}$$

Calculate weight of sodium lactate:

$$\text{Sodium lactate (g/2.5L)} = \frac{[Lact\text{-}] \text{ mmol/L} \times MW \times 2.5}{1000}$$

MW $CH_3CH(OH)COONa = 23 + (3 \times 12) + (5 \times 1) + (3 \times 16) = 112$

$$\text{Wt sodium lactate} = \frac{142 \times 112 \times 2.5}{1000} = \mathbf{39.8\ g}$$

Calculate weight of lactic acid:

$$\text{Lactic acid (g/2.5L)} = \frac{[LactH] \text{ mmol/L} \times MW \times 2.5}{1000}$$

MW $CH_3CH(OH)COOH = (3 \times 12) + (6 \times 1) + (3 \times 16) = 90$

$$\text{Wt lactic acid} = \frac{1 \times 90 \times 2.5}{1000} = 0.225 \text{ g}$$

Convert to mL 85% lactic acid SG = 1.2:

$$\text{Vol lactic acid(mL)} = \frac{\text{Wt (g)} \times 100}{\text{\% purity} \times \text{SG}}$$

$$= \frac{0.225 \times 100}{85 \times 1.2} = \mathbf{0.22\ mL} \text{ (2 sig figs)}$$

7. *A 24h urine collection has a pH of 5.5 and total phosphate content of 65 mmol. If the arterial pH is 7.40 and the pK_a for phosphate is 6.80, how many millimoles of hydrogen ion are excreted as titratable acidity using HPO_4^{2-} as buffer?*

The reaction occurring when secreted hydrogen ions are buffered by phosphate in the glomerular filtrate is:

$$HPO_4^{2-} + H^+ \rightarrow H_2PO_4^-$$

And the corresponding Henderson-Hasselbalch equation is:

$$pH = pKa + \log_{10} \frac{[HPO_4^{2-}]}{[H_2PO_4^-]}$$

Calculate the ratio of the two phosphate ions in fresh glomerular filtrate (i.e. pH = 7.4):

$$7.4 = 6.8 + \log_{10} \frac{[HPO_4^{2-}]}{[H_2PO_4^-]}$$

$$\log_{10} \frac{[HPO_4^{2-}]}{[H_2PO_4^-]} = 7.4 - 6.8 = 0.6$$

$$\frac{[HPO_4^{2-}]}{[H_2PO_4^-]} = \text{antilog}_{10}\, 0.6 = 3.98 \ldots\ldots\ldots\text{(i)}$$

For simplicity assume a urine volume of 1 L so that the total phosphate concentration is 65 mmol/L. (We are told the amount not the concentration but since we are dealing with ratios and volume could be used).

$$65 = [HPO_4^{2-}] + [H_2PO_4^-]$$

$$[H_2PO_4^-] = 65 - [HPO_4^{2-}]$$

Substitute for $[H_2PO_4^-]$ in (i) then solve for $[HPO_4^{2-}]$:

$$\frac{[HPO_4^{2-}]}{65 - [HPO_4^{2-}]} = 3.98$$

$$[HPO_4^{2-}] = 3.98\,(65 - [HPO_4^{2-}])$$

$$[HPO_4^{2-}] + 3.98\,[HPO_4^{2-}] = 258.7$$

$$4.98\,[HPO_4^{2-}] = 258.7$$

$$[HPO_4^{2-}] = \frac{258.7}{4.98} = 51.9 \text{ mmol/L} \quad (3 \text{ sig figs})$$

Repeat this procedure for acidified glomerular filtrate i.e. urine pH = 5.5

$$5.5 = 6.8 + \log_{10} \frac{[HPO_4^{2-}]}{[H_2PO_4^-]}$$

$$\log_{10} \frac{[HPO_4^{2-}]}{[H_2PO_4^-]} = 5.5 - 6.8 = -1.3$$

$$\frac{[HPO_4^{2-}]}{[H_2PO_4^-]} = \text{antilog}_{10}\ -1.3 = 0.050$$

Substitute $[H_2PO_4^-] = 65 - [HPO_4^{2-}]$ and solve for $[HPO_4^{2-}]$:

$$\frac{[HPO_4^{2-}]}{65 - [HPO_4^{2-}]} = 0.050$$

$$[HPO_4^{2-}] = 0.050\,(65 - [HPO_4^{2-}])$$

$$[HPO_4^{2-}] + 0.050[HPO_4^{2-}] = 0.050 \times 65 = 3.25$$

$$1.05\,[HPO_4^{2-}] = 3.25$$

$$[HPO_4^{2-}] = \frac{3.25}{1.05} = 3.10 \text{ mmol/L}$$

The titratable acidity is the concentration (or rather amount) of HPO_4^{2-} consumed:

Titratable acidity = 51.9 - 3.1 = **49 mmol** (2 sig figs)

8. *A buffer solution (pH 4.74) contains acetic acid (0.1 mol/L) and sodium acetate (0.1 mol/L) i.e. it is a 0.2M acetate buffer. Calculate the pH after addition of 4 mL of 0.025 M hydrochloric acid to 10 mL of the buffer.*

The dissociation to be considered is:

$$HAc \rightarrow H^+ + Ac^-$$

And the relevant form of the Henderson Hasselbalch equation is:

$$pH = pKa + \log_{10} \frac{[Ac^-]}{[HAc]}$$

Determine pKa by substituting the pH (4.74) and concentrations of Ac^- (0.1 mol/L) and HAc (0.1 mol/L):

$$4.74 = pKa + \log_{10} \frac{0.1}{0.1}$$

Since 0.1/0.1 = 1 and $\log_{10}$ 1 is 0, then pKa = 4.74

Calculate the adjusted cocnetrations of Ac^- and HAc, and substitute into the Henderson-Haseelbalch equation (using pKa = 4.74) then solve for pH:

Addition of HCl to this buffer converts some of the acetate ions to acetic acid:

$$Ac^- + H^+ \rightarrow HAc$$

Final $[Ac^-]$ = Initial $[Ac^-]$ - Added [HCl]

Allowance must be made for the dilution resulting from mixing 10 mL buffer with 4 mL HCl (total volume = 14 mL):

$$\text{Initial } [Ac^-] = \frac{0.1 \times 10}{14} = 0.071 \text{ mol/L}$$

$$\text{Added [HCl]} = \frac{\text{Initial [HCl]} \times 4}{14} = \frac{0.025 \times 4}{14} = 0.0071 \text{ mol/L}$$

Final $[Ac^-]$ = 0.071 - 0.0071 = 0.0639 mol/L

Similarly:

Final [HAc] = Initial [HAc] + Added [HCl]

Since Initial [HAc] = Initial $[Ac^-]$

Final [HAc] = 0.071 + 0.0071 = 0.0781 mol/L

Therefore: $pH = 4.74 + \log_{10} \frac{0.0639}{0.0781}$

$= 4.74 + \log_{10} 0.818$

$= 4.74 + (-0.087) =$ **4.65**

Chapter 4

1. *An aqueous solution in a 1 cm cell has an absorbance of 0.23 when read against a water blank at 500 nm. Assuming Beer's Law is obeyed, what volume of this solution would need to be added to 100 mL of water to give a solution which absorbs 30% of the light entering it under the same measurement conditions?*

   Only absorbance is proportional to concentration, so the first step is to calculate the absorbance of the final solution which absorbs 30% of the light entering into it.

   $$\text{Absorbance } (A) \quad = \quad \log_{10} \frac{I_0}{I}$$

   I_o = intensity of incident light = 100%

   I = intensity of transmitted light. Since 30% was absorbed, 100 – 30 = 70% is transmitted. Therefore, I = 70%

   Substitute these values to obtain A:

   $$A = \log_{10} \frac{100}{70} = \log_{10} 1.429 = 0.155 \text{ (3 sig figs)}$$

   Let x = vol of solution to be added to 100 mL water = x mL

   Final volume = $(100 + x)$ mL

   $$x \text{ (mL)} \times 0.23 = (100 + x) \times 0.155$$

   $$x \text{ (mL)} = \frac{(100 + x) \times 0.155}{0.23}$$

   $$x = \frac{15.5 + 0.155x}{0.23}$$

   $$0.23x = 15.5 + 0.155x$$

$$0.23\,x - 0.155\,x = 15.5$$

$$0.075\,x = 15.5$$

$$x = \frac{15.5}{0.075} = \mathbf{207\ mL} \quad (3 \text{ sig figs})$$

2. *Calculate the absorbances corresponding to the following percentage transmittance readings:*

a) 95 b) 75 c) 50 d) 25 e) 10 f) 1

If I_o is the intensity of incident light and I the intensity of transmitted light, then:

$$\text{transmittance } (\%T) = \frac{I \times 100}{I_o} \quad \text{and} \quad \text{absorbance } (A) = \log_{10}\frac{I_o}{I}$$

The expression for $\%T$ can be arranged to: $\dfrac{I_o}{I} = \dfrac{100}{\%T}$

Which can be substituted into the expression for A to give:

$$A = \log_{10}\frac{100}{\%T} = \log_{10} 100 - \log_{10} \%T$$

Substituting 2 for $\log_{10} 100$ gives the following useful expression:

$$\boldsymbol{A = 2 - \log_{10} \%T}$$

All that is required is to substitute values for $\%T$ into this expression to obtain A:

a) *%T = 95*

$A = 2 - \log_{10} 95 = 2 - 1.978 =$ **0.022** (3 sig figs)

b) *%T = 75*

$A = 2 - \log_{10} 75 = 2 - 1.875 =$ **0.125** (3 sig figs)

c) *%T = 50*

$A = 2 - \log_{10} 50 = 2 - 1.699 =$ **0.301** (3 sig figs)

d) *%T = 25*

$A = 2 - \log_{10} 25 = 2 - 1.398 =$ **0.602** (3 sig figs)

e) *%T = 10*

$A = 2 - \log_{10} 10 = 2 - 1.000 =$ **1**

f) *%T = 1*

$A = 2 - \log_{10} 1 = 2 - 0 =$ **2**

3. Calculate the % of incident light transmitted by solutions with the following absorbances:

a) 0.1 b) 0.25 c) 0.50 d) 0.75 e) 1.00 f) 2.00

The expression used to calculate A from $\%T$ can be rearranged to enable direct calculation of $\%T$ from A:

$$A = 2 - \log_{10} \%T$$

$$A + \log_{10} \%T = 2$$

$$\log_{10} \%T = 2 - A$$

$$\mathbf{\%T = antilog_{10} (2 - A)}$$

Therefore substitute values for A into this expression then evaluate $\%T$:

a) $A = 0.1$

$\%T = \text{antilog}_{10} (2 - 0.1) = \text{antilog}_{10} 1.9 =$ **79%** (2 sig figs)

b) $A = 0.25$

$\%T = \text{antilog}_{10} (2 - 0.25) = \text{antilog}_{10} 1.75 =$ **56%** (2 sig figs)

c) $A = 0.50$

$\%T = \text{antilog}_{10} (2 - 0.50) = \text{antilog}_{10} 1.50 =$ **32%** (2 sig figs)

d) $A = 0.75$

$\%T = \text{antilog}_{10} (2 - 0.75) = \text{antilog}_{10} 1.25 =$ **18%** (2 sig figs)

e) $A = 1.00$

$\%T = \text{antilog}_{10} (2 - 1.00) = \text{antilog}_{10} 1 =$ **10%**

f) $A = 2.00$

$\%T = \text{antilog}_{10} (2 - 2) = \text{antilog}_{10} 0 =$ **1%**

4. *A solution of a compound (concentration 100 mmol/L) was placed in a cuvette with a 1 cm light path and the percentage of incident light transmitted was 18.4. Calculate the molar absorptivity of the compound.*

First convert percentage transmittance (%T) to absorbance (A):

$$A = 2 - \log_{10} \%T$$

$$= 2 - \log_{10} 18.4$$

$$= 2 - 1.2648$$

$$= 0.735 \quad \text{(3 sig figs)}$$

Use this absorbance to calculate molar absorptivitiy using the Beer-Lambert Law:

$$A = abc$$

Where A = absorbance reading = 0.735
a = molar absorptivity = unknown
b = light path length = 1 cm
c = concentration = 100 mmol/L

Since the question asks for calculation of molar absorptivity, the concentration must be divided by 1,000 to convert it to mol/L (1 mol = 1,000 mmol):

$$\text{Concentration (mol/L)} = \frac{100 \text{ (mmol/L)}}{1{,}000} = 0.1 \text{ mol/L}$$

Substitute A, b and c into the Beer-Lambert equation and solve for a:

$$0.735 = a \times 1 \times 0.1$$

$$a = \frac{0.735}{1 \times 0.1} = 7.35$$

The units can be derived by entering the individual units into the same equation (remembering that absorbance is the logarithm of a ratio so has no units):

$$a = \frac{-}{\text{cm x mol/L}} = \frac{\text{L}}{\text{cm x mol}} = \text{L/cm/mol (or L.mol}^{-1}\text{.cm}^{-1}\text{)}$$

Therefore molar absorptivity = **7.35 L.mol^{-1}.cm^{-1}**

5. *The transmittance of a solution of NADH at 340 nm is 45%. What is the absorbance at 340 nm of a 1 in 5 dilution of this solution?*

Since absorbance, not transmittance, is linearly proportional to concentration, the first step is to convert the transmittance (%*T*) to absorbance (*A*):

$$A = 2 - \log_{10} \%T$$

$$= 2 - \log_{10} 45$$

$$= 2 - 1.6532$$

$$= 0.3468$$

Assuming NADH obeys Beer's Law, the absorbance of a 1 in 5 dilution of this solution will be a fifth of this value:

$$\text{Absorbance of 1 in 5 dilution} = \frac{0.3468}{5} = \mathbf{0.069} \text{ (2 sig figs)}$$

6. *75 mg of faeces were homogenised in 1 mL of concentrated hydrochloric acid, then 3 mL diethylether added, mixed, 3 mL of water added and mixed again. After centrifugation the aqueous phase (volume 4.5 mL) was scanned in a spectrophotometer using a cell with a 1 cm pathlength and the peak height at 405 nm due to porphyrin, after applying a background correction, was 0.35 absorbance units. A separate 0.250 g portion of faeces was dried in a 100°C oven for 3 hours after which it's weight was 0.125 g. Given that the molar absorption coefficient of porphyrin is 2.75 x 10^5 L/mol/cm calculate the porphyrin concentration in nmol/g dry weight of faeces.*

First use the Beer-Lambert equation to calculate the porphyrin concentration in the extract:

$$A = abc$$

Where A = absorbance = 0.35
a = molar absorptivity = 2.75 x 10^5 L/mol/cm
b = path length = 1 cm
c = concentration = mol/L

$$0.35 = 2.75 \times 10^5 \times 1 \times c$$

$$c = \frac{0.35}{2.75 \times 10^5 \times 1} = 1.273 \times 10^{-6} \text{ mol/L}$$

The answer is required in nmol not mol so this value must be multiplied by 10^9 (since 1 mol = 10^9 nmol):

$$c \text{ (nmol/L)} = 1.273 \times 10^{-6} \times 10^9 = 1.273 \times 10^3 \text{ nmol/L}$$

Since the faecal sample produced 4.5 mL of extract, the amount of porphyrin in the sample is obtained by dividing by 1,000 (to convert from nmol/L to nmol/mL), the multiplying by the volume of extract (4.5 mL):

$$\text{Porphyrin in sample} = \frac{1.273 \times 10^3 \times 4.5}{1{,}000 \ (= 10^3)} = 5.729 \text{ nmol}$$

The porphyrin content (expressed as nmol/g fresh weight of faeces) is obtained by dividing the porphyrin extracted (5.729 nmol) by the weight of sample used to prepare the extract (75 mg = 0.075 g):

$$\text{Porphyrin (nmol/g fresh stool)} = \frac{5.729}{0.075} = 76.4 \text{ nmol/g fresh wt} \quad (3 \text{ sig figs})$$

To express this result on a dry weight basis, multiply by the fresh weight of faeces used for the dry weight determination then divide by its dry weight:

$$\text{Porphyrin content} = \frac{76.4 \times 0.250}{0.125} = \textbf{153 nmol/g dry faeces (3 sig figs)}$$

7. *A solution containing a substance of molecular weight 400 at a concentration of 3 g/L transmitted 75% of incident light of a particular wavelength in a 1 cm cuvette. Calculate the % of incident light of the same wavelength that would be transmitted by a solution of the same substance at a concentration of 4 g/L and calculate the molar absorption coefficient for that substance at this wavelength.*

First convert the % transmittance (T) to absorbance (A):

$$A = 2 - \log_{10} \%T$$

$$A = 2 - \log_{10} 75$$

$$= 2 - 1.875$$

$$= 0.125$$

Provided the substance obeys Beer's Law over the range of concentrations (i.e. absorbance is directly proportional to concentration), then the absorbance of a different concentration (4 g/L) can be calculated from the relationship:

$$\frac{\text{Absorbance}_2}{\text{Concentration}_2} = \frac{\text{Absorbance}_1}{\text{concentration}_1}$$

Rearranging: $\text{Absorbance}_2 = \frac{\text{Absorbance}_1 \times \text{Concentration}_2}{\text{Concentration}_1}$

Substitute: absorbance_2 = unknown ; absorbance_1 = 0.125;
concentration_2 = 4 g/L; concentration_1 = 3 g/L

$$\text{Absorbance}_{(4\text{g/L})} = \frac{4 \times 0.125}{3} = 0.1667 \text{ (4 sig figs)}$$

Convert this absorbance to % transmittance:

$$\text{Log}_{10}\%T = 2 - A = 2 - 0.1667 = 1.833$$

$$\%T = \text{antilog}_{10}\, 1.833 = \mathbf{68\%} \text{ (2 sig figs)}$$

To calculate molar absorption coefficient use either pairs of concentration and absorption:

$$A = abc$$

Where A = absorbance = 0.167 (for a concentration of 4 g/L)
a = molar absorption coefficient = ?
b = cell path length = 1 cm
c = concentration = 4g/L.
Since MW = 400
molar concentration = 4/400 = 1/100 = 1.0×10^{-2} mol/L

$$0.167 = a \times 1.0 \times 1.0 \times 10^{-2}$$

$$a = \frac{0.167}{1.0 \times 10^{-2}} = 0.167 \times 10^{2} = \mathbf{16.7\ L.mol^{-1}.cm^{-1}}$$

8. *The absorbances of a solution containing NAD and NADH in a 1cm light path cuvette were 0.337 at 340 nm and 1.23 at 260 nm. The molar extinction coefficients are:*

NAD: 1.8×10^4 at 260 nm, 1.0×10^{-3} at 340 nm
NADH: 1.5×10^4 at 260 nm, 6.3×10^3 at 340 nm

Calculate the concentrations of NAD and NADH in the solution.

Both NAD and NADH absorb at the two wavelengths used (260 nm and 340 nm). Absorbances are additive, therefore at either wavelength:

Total absorbance = Absorbance of NAD + Absorbance of NADH

At any wavelength the absorbance of NAD or NADH is given by:

Absorbance = Molar extinction coefficient x Molar concentration x Cell path

Therefore for each wavelength equations can be set up relating measured total absorbance to the sums of the individual absorbances of NAD and NADH:

Measured absorbance = (NAD_{Conc} x NAD_{Coeff}) + ($NADH_{Conc}$ x $NADH_{Coeff}$)

At 340 nm: $0.337 = 1.0 \times 10^{-3}$ [NAD] + 6.3×10^3 [NADH] ..... (i)

At 260 nm: $1.23 = 1.8 \times 10^4$ [NAD] + 1.5×10^4 [NADH] ..... (ii)

(The cell path is 1 cm and can be ignored)

These form a pair of simultaneous equations which can be solved for [NAD] and [NADH] in the usual manner. However, solving a set of simultaneous equations can be a lengthy process. Therefore we should look for approximations and short cuts. In this particular example it is possible to considerably simplify the calculation. The molar extinction coefficient of NAD at 340 nm is much lower than that of NADH (by a factor of approx. 10^{-6}) so that the contribution of NAD to absorbance at this wavelength can be ignored. Equation (i) can then be simplified to:

$0.337 = 6.3 \times 10^3$ [NADH]

$$[NADH] = \frac{0.337}{6.3 \times 10^3} = 5.35 \times 10^{-5} \text{ M} = \textbf{53.5 µmol/L}$$

[NAD] can be calculated by substituting [NADH] = 5.35×10^{-5} into equation (ii):

$1.23 = 1.8 \times 10^4$ [NAD] + $(1.5 \times 10^4 \times 5.35 \times 10^{-5})$

$1.23 = 1.8 \times 10^4$ [NAD] + (8.03×10^{-1})

1.8×10^4 [NAD] = $1.23 - (8.03 \times 10^{-1}) = 0.427$

$$[NAD] = \frac{0.427}{1.8 \times 10^4} = 2.37 \times 10^{-5} \text{ M} = \textbf{23.7 µmol/L}$$

9. *25 mg of bilirubin ($C_{33}H_{36}O_6N_4$) were dissolved in 4 mL of dimethyl sulphoxide; 200 µL of this solution was diluted to 250 mL with chloroform. This solution gave an absorbance of 0.502 when measured in a 1 cm cell against a chloroform blank.*

Given that the molar absorptivity of bilirubin under these conditions is 6.07×10^4, calculate the percentage purity of the bilirubin.

First calculate the concentration of bilirubin in the final solution:

$$A = a \times b \times c$$

Where
A = absorbance = 0.502
a = molar absorbivity = 6.07×10^4 L.mol^{-1}.cm^{-1}
b = path length = 1 cm
c = concentration in mol/L = ?
1 = path length = 1 cm

$0.502 = 6.07 \times 10^4 \times 1 \times c$

Rearranging and solving for *c*:

$$c = \frac{0.502}{6.07 \times 10^4} = 8.27 \times 10^{-6} \text{ mol/L} = 8.27 \times 10^{-3} \text{ mmol/L}$$

Use this concentration of the final solution to calculate the bilirubin content of the weighed bilirubin:

The final solution was prepared by diluting 200 μL (i.e. 0.2 mL) of stock to 250 mL

Therefore actual concentration of stock

$$= \frac{8.27 \times 10^{-3} \times 250}{0.2} = 10.34 \text{ mmol/L}$$

4 mL (the volume of DMSO the bilirubin was dissolved in) contains:

$$\frac{10.34 \times 4}{1000} = 0.04136 \text{ mmol bilirubin}$$

Convert to wt of bilirubin:

$$\text{Wt bilirubin (mg)} = \text{mmol bilirubin} \times \text{MW}$$

$$\text{MW bilirubin} = (33 \times 12) + (36 \times 1) + (6 \times 16) + (4 \times 14) = 584$$

$$\text{Therefore, wt bilirubin} = 0.0414 \times 584 = 24.15 \text{ mg}$$

$$\% \text{ purity} = \frac{\text{Amount of bilirubin by assay} \times 100}{\text{Amount of bilirubin weighed}}$$

$$= \frac{24.15 \times 100}{25} = \mathbf{97}\% \text{ (2 sig figs)}$$

10. *A method for creatinine determination based on the Jaffe reaction involved mixing 0.1 mL of sample with 2.5 mL alkaline picrate reagent, incubating for 10 min at room temperature, then measuring the absorbance at 530 nm in a 1-cm cuvette in a spectrophotometer set to read zero using a cuvette containing distilled water. The following readings were obtained:*

Blank (water as ample)	*0.050*
Creatinine standard (200 µmol/L)	*0.250*
Serum sample	*0.125*
Urine sample (prediluted 1 in 50 with water)	*0.200*

Calculate the creatinine concentration in the serum (in µmol/L) and urine (in mmol/L).

First subtract the reagent blank (i.e. the reading obtained when using water as sample) from each absorbance reading:

	Absorbance	Corrected Absorbance
Blank (water as ample)	0.050	0.000
Creatinine standard (200 µmol/L)	0.250	0.200
Serum sample	0.125	0.075
Urine sample (prediluted 1 in 50 with water)	0.200	0.150

$$\frac{\text{Corrected absorbance of unknown}}{\text{Concentration of unknown}} = \frac{\text{Corrected absorbance of standard}}{\text{Concentration of standard}}$$

$$\text{Concentration of unknown} = \frac{\text{Corrected absorbance of unknown} \times \text{Concentration of standard}}{\text{Corrected absorbance of standard}}$$

For serum:

$$\text{Serum creatinine (µmol/L)} = \frac{0.075 \times 200}{0.200} = \mathbf{75\ \mu mol/L}$$

For urine the calculation is the same except that the result must be multiplied by 50 to allow for the predilution of the sample prior to assay, then divided by 1,000 to convert from µmol/L to mmol/L (1 mmol = 1,000 µmol):

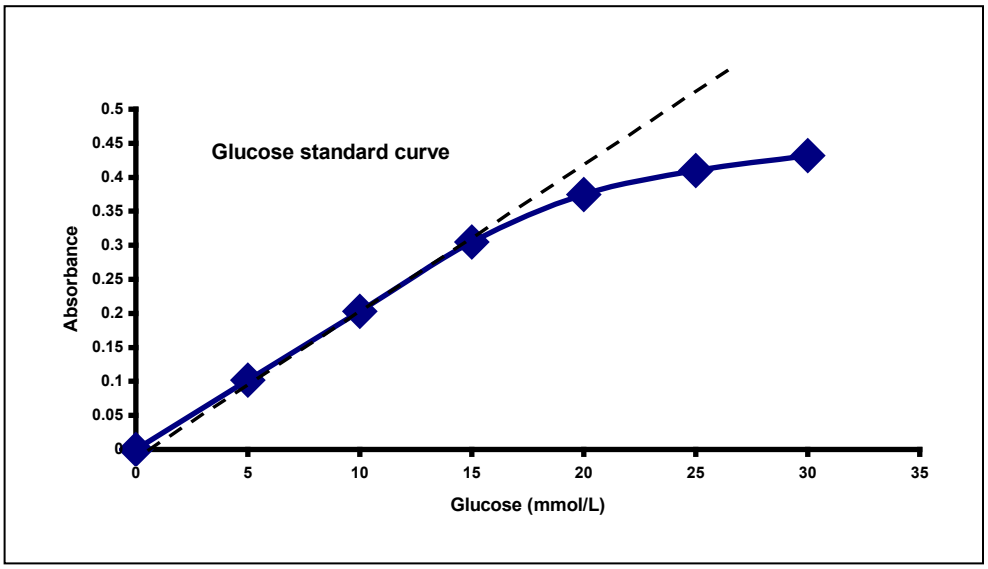
$$\text{Urine creatinine (mmol/L)} = \frac{0.150 \times 50 \times 200}{0.200 \times 1{,}000} = \textbf{7.5 mmol/L}$$

11. *A standard curve for a plasma glucose method was set up by preparing a series of dilutions of a stock glucose standard (containing 50 mmol glucose/L) and measuring the absorbance at 500 nm in a 1 cm cuvette using a blank with zero glucose concentration to zero the instrument. The following readings were obtained:*

Glucose (mmol.L):	*5*	*10*	*15*	*20*	*25*	*30*
Absorbance:	*0.102*	*0.203*	*0.305*	*0.375*	*0.410*	*0.432*

Does the method obey Beer's Law? What glucose concentration corresponds to an absorbance reading of 0.250?

Plot the absorbance (vertical scale) against the standard concentration (horizontal scale) including the zero as a point (since the blank was used to zero the instrument):



Inspection of the curve shows that the method only obeys Beer's Law up to a concentration of 15 mmol/L (when absorbance = 0305).

The slope of the curve up to this point $= \frac{0.305}{15} = 0.020\ A/\text{mmol}$

Therefore $1\ A = \frac{1}{0.02}$ so that $0.250\ A = \frac{0.250}{0.02} =$ **12.5 mmol/L**

Chapter 5

1. *An aliquot of a 24 h urine (volume 1850 mL) has a creatinine concentration of 8500 µmol/L. Calculate the 24 h urinary creatinine excretion expressing the result as mmol/24 h.*

Creatinine excretion (mmol/24h) =

creatinine concentration (mmol/L) x 24 h urine volume (L)

Divide the creatinine concentration by 1,000 to covert from µmol/L to mmol/L (1,000 µmol = 1 mmol).

Divide the urine volume by 1,000 to convert fro mL to l L (1,000 mL = 1 L).

$$\text{Creatinine excretion (mmol/24 h)} = \frac{8{,}500 \times 1{,}850}{1{,}000 \times 1{,}000} = \textbf{15.7 mmol/24 h}$$

2. *A patient has a GFR of 110 mL/min. If the plasma creatinine concentration is 180 µmol/L how many mmol of creatinine are filtered in 12 h?*

First convert the filtration rate from mL/min to L/12 h. Multiply by 60 (to convert from min to h), then by 12 (to convert from h to 12 h) and finally divide by 1,000 (to convert from mL to L):

$$\text{Filtration rate (L/12 h)} = \frac{110 \times 60 \times 12}{1{,}000} = 79.2 \text{ L/12h}$$

The answer is required in mmol so divide the plasma concrentration by 1,000 to convert from µmol/L to mmol/L (1,000 µmol = 1 mmol):

$$\text{Plasma creatinine (mmol/L)} = \frac{180}{1{,}000} = 0.18 \text{ mmol/L}$$

Creatinine filtered (mmol/12 h) =

Plasma creatinine (mmol/L) x Filtration rate (L/12 h)

Creatinine filtered (mmol/12 h) = 0.18 x 79.2 = **14.3 mmol/12 h**

3. *A urine collection (volume 3.2 L) was handed in by a patient which he said he had collected over the previous day. Calculate the creatinine clearance given that the urine was found to have a creatinine concentration of 7.2 mmol/L. The plasma creatinine concentration taken during the collection was 94 µmol/L. Give the most likely cause for this result.*

Creatinine clearance(mL/min) =

$$\frac{\text{Urine creatinine (mmol/L)} \times \text{Urine flow rate (mL/min)}}{\text{Plasma creatinine (mmol/L)}}$$

Urine creatinine concentration = 7.2 mmol/L

Plasma creatinine = 94 µmol/L = $\frac{94}{1{,}000}$ mmol/L

Urine flow rate = 3.2 L/24 h = $\frac{3.2}{24}$ L/h = $\frac{3.2}{24 \times 60}$ L/min = $\frac{3.2 \times 1{,}000}{24 \times 60}$ mL/min

Creatinine clearance = $\frac{7.2 \times 3.2 \times 1{,}000 \times 1{,}000}{24 \times 60 \times 94}$ = **170 mL/min**

This creatinine clearance is a little high. The most likely cause is that the 24 collection was made over a longer period than 24 h – perhaps the bladder was not emptied at the start of the collection period (or if emptied it was added to the collection instead of being discarded).

4. *The concentration of a compound in the plasma of a normal adult is 10 mg/L. The GFR is 110 mL/min and 316.8 mg of the compound are excreted over 24 h in a urine volume of 1584 mL. Comment on these findings.*

First calculate the total amount of the compound filtered over a 24 h period (based on the assumption that the compound is freely filtered at the glomerulus):

$$\text{Amount filtered (mg)} = \text{GFR (L/24h)} \times \text{Plasma concentration (mg/L)}$$

$$\text{GFR} = 110 \text{ mL/min} = \frac{110}{1{,}000} \text{ L/min} = \frac{110 \times 60}{1{,}000} \text{ L/h} = \frac{110 \times 60 \times 24}{1{,}000} \text{ L/24 h}$$

$$\text{Amount filtered (mg/24 h)} = \frac{110 \times 60 \times 24 \times 10}{1{,}000} = \mathbf{1584 \text{ mg/24 h}}$$

The rate of excretion (316.8 mg/24 h) is much less than this suggesting either that the compound is either not freely filtered at the glomerulus or considerable amounts are reabsorbed from the filtrate.

N.B. The urine volume was not used in this calculation. Another approach (which would utilize urine volume) would be to calculate the clearance of the compound (which comes out at 22 mL/min) then compare it with the GFR.

5. *A subject with a GFR of 100 mL/min was infused with a 'drug' X at a rate of 100 µmol/min and the plasma concentration reached a steady state value of 200 µmol/L. It is known that this drug is not metabolized or excreted by organs other than the kidney. What is the clearance of this drug? Comment on the result.*

When a steady state is reached the rate of excretion is equal to the rate of infusion and the plasma concentration reaches a constant value.

$$\text{Clearance (mL/min)} = \frac{\text{Excretion rate (µmol/min)}}{\text{Plasma concentration (µmol/mL)}}$$

Excretion rate = infusion rate = 100 µmol/min

Plasma concentration = 200 µmol/L = $\frac{200}{1{,}000}$ µmol/mL

Clearance (mL/min) = $\frac{100 \text{ x } 1000}{200}$ = **500 mL/min**

The clearance of the drug far exceeds the GFR suggesting that the mode of excretion is predominantly tubular secretion.

6. *A patient who is severely water depleted and excreted only 100 mL of urine in the last 6 hours was a short time before, found to have a creatinine clearance of 100 mL/min with a plasma creatinine concentration of 100 µmol/L. If renal function has remained unchanged what concentration of creatinine would you expect to find in the latest 100 mL (6 h collection) specimen of urine?*

This question involves calculating the urinary excretion when the plasma concentration and clearance is known. The expression for clearance is:

$$\text{Clearance (mL/min)} = \frac{\text{Urine creatinine (mmol/L)} \text{ x } \text{Urine flow rate (mL/min)}}{\text{Plasma creatinine (mmol/L)}}$$

Rearranging this expression:

$$\text{Urine creatinine (mmol/L)} = \frac{\text{Clearance (mL/min)} \text{ x } \text{Plasma creatinine (mmol/L)}}{\text{Urine flow rate (mL/min)}}$$

Creatinine clearance = 100 mL/min

Plasma creatinine = 100 µmol/L = $\frac{100}{1{,}000}$ mmol/L

Urine flow rate = 100 mL/6 h = $\frac{100}{6}$ mL/h = $\frac{100}{6 \text{ x } 60}$ mL/min

$$\text{Urine creatinine (mmol/L)} = \frac{100 \times 100 \times 6 \times 60}{1{,}000 \times 100} = \textbf{36 mmol/L}$$

7. *Estimate the effect on urinary sodium excretion in a person with a GFR of 95 mL/min and plasma sodium of 140 mmol/L, of a 1% decrease in the overall reabsorption of sodium.*

First calculate the amount of sodium filtered at the glomerulus in mmol/min:

$$\text{Na filtered (mmol/min)} = \text{GFR (mL/min)} \times \text{Plasma Na (mmol/mL)}$$

$$\text{Plasma Na} = 140 \text{ mmol/L} = \frac{140}{1{,}000} \text{ mmol/mL}$$

$$\text{Na filtered (mmol/min)} = \frac{95 \times 140}{1{,}000} = 13.3 \text{ mmol/min}$$

If the amount reabsorbed decreases by 1% then the amount excreted in the urine will increase by 1% of that filtered:

$$\text{Increase in Na excretion} = \frac{13.3 \times 1}{100} = 0.133 \text{ mmol/min}$$

Therefore the increase in urine Na over a 24 h period =

$$0.133 \times 60 \times 24 = \textbf{192 mmol/24 h}$$

8. *The following data were obtained for a hypertensive patient on a low sodium diet:*

Plasma: *creatinine = 200 µmol/L* *sodium = 155 mmol/L*
24 h Urine: *creatinine = 12.5 mmol/L* *volume = 1250 mL*

If the renal tubules reabsorb 90% of filtered sodium, how many grams of sodium are excreted in the same 24 h period?

$$\text{Fractional excretion}_{Na} = \frac{\text{Na excreted in urine}}{\text{Na filtered}}$$

If 90% of filtered Na is reabsorbed then 100 – 90 = 10% must excreted i.e. fractional excretion (FE_{Na}) = 10%

FE_{Na} is calculated from the expression:

$$FE_{Na}\ (\%) = \frac{(\text{Urine}_{Na} \times \text{Plasma}_{Creatinine}) \times 100}{\text{Urine}_{Creatinine} \times \text{Plasma}_{Na}}$$

Which can be re-arranged to give an expression for urine sodium:

$$\text{Urine}_{Na} = \frac{FE_{Na} \times \text{Urine}_{Creatinine} \times \text{Plasma}_{Na}\ \text{mmol/L}}{\text{Plasma}_{Creatinine} \times 100}$$

Substitute these values to obtain the urine sodium concentration N.B units must be the same so convert plasma creatinine (µmol/L) to mmol/L by dividing it by 1,000.

$$\text{Urine}_{Na} = \frac{10 \times 12.5 \times 155 \times 1{,}000}{200 \times 100} = 969\ \text{mmol/L}$$

Since the 24 h urine volume is 1250 mL (= 1.25 L) the amount excreted in 24 h is:

$$969 \times 1.25 = 1211\ \text{mmol/24 h}$$

Convert to g/24 h:

$$\text{Na (g/24 h)} = \text{Na (mol/24 h)} \times \text{MW}$$

Divide the sodium output by 1,000 to convert from mmol/24 to mol/24 h. MW Na = 23.

$$\text{Na (g/24 h)} = \frac{1211 \times 23}{1{,}000} = \textbf{28 g} \text{ (2 sig figs)}$$

Or expressed as NaCl (MW = 58.5):

$$\text{NaCl (g/24 h)} = \frac{1211 \times 58.5}{1{,}000} = \textbf{71 g} \text{ (2 sig figs)}$$

Another approach to this problem would be to first calculate the GFR from the creatinine results, then use this to calculate the Na filtered etc.

9. *The following results were obtained in a 20 year old male admitted after a car crash and found to be oliguric:*

Plasma	*Na*	*135 mmol/L*
	K	*5.0 mmol/L*
	Urea	*25.0 mmol/L*
	Creatinine	*250 μmol/l*
Urine	*Na*	*90 mmol/L*
	Creatinine	*2.4 mmol/L*
	Osmolality	*200 mOsm/kg*

Calculate the fractional excretion of sodium.

$$FE_{Na} = \frac{\text{Urine}_{Na} \times \text{Plasma}_{Creatinine}}{U_{Creatinine} \times \text{Plasma}_{Na}}$$

All units need to be the same, if mmol/L used then:

$$\text{Plasma creatinine} = 250\ \mu\text{mol/L} = \frac{250}{1{,}000}\ \text{mmol/L}$$

$$FE_{Na} = \frac{90 \times 250}{1{,}000 \times 2.4 \times 135} = \textbf{0.069} \text{ (or 6.9\%)}$$

10. A 45 year old lady has a body weight of 56 kg and a height of 155 cm. If her plasma creatinine is 150 µmol/L estimate her GFR expressing the result as mL/min/1.73 m^2.

Since body weight is given, the Cockcroft-Gault formula for females can be used:

$$\text{GFR (mL/min)} = \frac{(140 - \text{age in yrs}) \times \text{Body wt (kg)} \times 1.2 \times 0.85}{\text{Plasma creatinine (µmol/L)}}$$

$$\text{GFR (mL/min)} = \frac{(140 - 45) \times 56 \times 1.2 \times 0.85}{150}$$

$$= 36 \text{ mL/min}$$

Next calculate the patient's body surface area (A) using the body weight in kg (W) and height in cm (H):

$$A = \text{antilog}_{10}[(0.425 \times \log_{10} W) + (0.725 \times \log_{10} H) - 2.144] \text{ m}^2$$

$$A = \text{antilog}_{10}[(0.425 \times \log_{10} 56) + (0.725 \times \log_{10} 155) - 2.144]$$

$$= \text{antilog}_{10}[(0.425 \times 1.75) + (0.725 \times 2.19) - 2.144]$$

$$= \text{antilog}_{10}[0.744 + 1.588 - 2.144]$$

$$= \text{antilog}_{10}\ 0.188$$

$$= 1.54 \text{ m}^2$$

$$\text{Corrected } GFR \text{ (mL/min/1.73m}^2) = \frac{\text{Measured } GFR \text{ (mL/min)} \times 1.73}{A \text{ (m}^2)}$$

$$= \frac{36 \times 1.73}{1.54}$$

$$= \mathbf{40 \text{ mL mL/min/1.73 m}^2}$$

Alternatively the abbreviated MDRD formula can be used (height not required).

11. Calculate the tubular maximum reabsorptive capacity (Tm/GFR) for glucose from the following data:

Plasma glucose	*10 mmol/L*	*Plasma creatinine*	*120 µmol/L*
Urine glucose	*50 mmol/L*	*Urine creatinine*	*6.0 mmol/L*

The urine (volume 30 mL) was collected over a 15 minute period.

First calculate the fractional excretion of glucose ($FE_{Glucose}$):

$$FE_{Glucose} = \frac{Urine_{Glucose} \times Plasma_{Creatinine}}{Urine_{Creatinine} \times Plasma_{Glucose}}$$

All units must be the same so first correct plasma creatinine to mmol/L:

$$\text{Plasma creatinine} = 120\ \mu\text{mol/L} = \frac{120}{1{,}000}\ \text{mmol/L}$$

$$FE_{Glucose} = \frac{50 \times 120}{1{,}000 \times 6.0 \times 10} = 0.1$$

This is the fraction of filtered glucose which is NOT reabsorbed by the tubules. The fraction reabsorbed (*TR*) is next calculated:

$$TR = 1 - FE = 1 - 0.1 = 0.9$$

To convert this reabsorption fraction to the absolute amount reabsorbed (i.e. The *Tm*/*GFR*), multiply by the plasma concentration:

$$Tm/GFR = TR \times \text{Plasma concentration}$$

$$= 0.9 \times 10$$

$$= \mathbf{9\ mmol/L\ glomerular\ filtrate}$$

12 *A 6 h urine collection (volume 800 mL) has an osmolality of 200 mOsm/kg. If the plasma osmolality is 260 mOsm/kg calculate the free water clearance in mL/min.*

First calculate the osmolar clearance (C_{osm}):

$$C_{osm} = \frac{U_{osm} \times V}{P_{osm}}$$

P_{osm} = 260 mOsm/kg
U_{osm} = 200 mOsm/kg
V = urine flow rate = 800 mL/6mh

$$= \frac{800}{6} \text{ mL/h} = \frac{800}{6 \times 60} \text{ mL/min} = 2.22 \text{ mL/min}$$

$$C_{osm} = \frac{2\,00 \times 2.22}{260} = 1.71 \text{ mL/min}$$

The free water clearance (C_{water}) is the difference between the urine flow rate and the osmolar clearance:

$$C_{water} = V - C_{osm}$$

$$C_{water} = 2.22 - 1.71 = \textbf{0.51 mL/min}$$

13. *An estimation of glomerular filtration rate can be calculated using the abbreviated MDRD (Modified Diet in Renal Disease) formula:*

$$GFR\ (mL/min/1.73m^2) = 186 \times [serum\ creatinine \times 0.011312]^{-1.154}$$

$$\times [age\ in\ years]^{-0.203} \times 0.742 \text{ if female and/or } \times 1.21 \text{ if Afro American origin}$$

(where serum creatinine is in µmol/L)

Calculate the GFR for a 57 year old Caucasian women whose serum creatinine is 130 µmol/L, and her creatinine clearance, given that a 24 h urine collection with a volume of 1.1 L had a creatinine concentration of 4.7 mmol/L.

Comment critically on the two values.

First calculate the GFR using the abbreviated MDRD formula by substituting values for serum creatinine (130 µmol/L) and age (57 y) – remembering to multiply by 0.742 since the patient is female:

$$\text{GFR (mL/min/1.73m}^2) = 186 \times [130 \times 0.011312]^{-1.154} \times [57]^{-0.203} \times 0.742$$

$$= 186 \times 1.471^{-1.154} \times 57^{-0.203} \times 0.74$$

$$= 186 \times \text{antilog}_{10}[-1.154 \times \log_{10} 1.471] \times \text{antilog}_{10}[-0.203 \times \log_{10} 57] \times 0.742$$

$$= 186 \times \text{antilog}_{10}[-1.154 \times 0.1676] \times \text{antilog}_{10}[-0.203 \times 1.7559] \times 0.742$$

$$= 186 \times \text{antilog}_{10}(-0.1934) \times \text{antilog}_{10}(-0.3565) \times 0.742$$

$$= 186 \times 0.6406 \times 0.4400 \times 0.742 = \mathbf{39\ mL/min/1.73\ m^2} \text{ (2 sig figs)}$$

Next calculate the creatinine clearance:

$$\text{Creatinine clearance (mL/min)} = \frac{\text{Urine creatinine (mmol/L)} \times \text{Urine flow rate (mL/min)}}{\text{Serum creatinine (mmol/L)}}$$

Urine creatinine = 4.7 mmol/L

$$\text{Serum creatinine} = 125\ \mu\text{mol/L} = \frac{130}{1{,}000}\ \text{mmol/L}$$

$$\text{Urine flow rate} = 1.1\ \text{L/24 h} = 1.1 \times 1{,}000\ \text{mL/24 h}$$

$$= \frac{1.1 \times 1{,}000}{24}\ \text{mL/h} = \frac{1.1 \times 1{,}000}{24 \times 60}\ \text{mL/min}$$

$$\text{Creatinine clearance (mL/min)} = \frac{4.7 \times 1.1 \times 1{,}000 \times 1{,}000}{24 \times 60 \times 130}$$

$$= \mathbf{28\ mL/min}$$

There are several possible reasons for the discrepancy between the derived GFR and the calculated clearance:

- Inaccuracy in the timed urine collection. **This is potentially the greatest source of error.** Although the 24 h volume of 1.1 L seems reasonable the calculated creatinine excretion seems low (1.1 x 4.7 = 5.2 mmol/24 h) - unless the lady has a very low muscle mass – suggesting that the collection is incomplete.

- Failure to correct the creatinine clearance for body surface area (this would require knowledge of weight and height). However, the MDRD formula does not take into account individual variation in body surface area either, but just assumes an average value based on the patient's age and sex.

- Creatinine is secreted by tubules into the urine so that creatinine clearance measurements always overestimates GFR.

Chapter 6

1. *Calculate the approximate osmolality of a glucose/saline infusion containing equal proportions of 5% glucose and 0.9% sodium chloride.*

First calculate the osmolalities due to glucose and sodium chloride individually.

Formula for glucose = $C_6H_{12}O_6$

AW C = 12, therefore C_6 = 6 x 12 = 72
AW H = 1, therefore H_{12} = 12 x 1 = 12
AW O = 16, therefore O_6 = 6 x 16 = 96

MW = 180

$$\text{Osmolality}_{\text{Glucose}} = \frac{\text{Glucose concentration (g/L)}}{\text{MW}}$$

Glucose concentration = initially 5% = finally 2.5% (since mixed with an equal volume of saline).

2.5% glucose = 2.5 g/100 mL = 25 g/L

$$\text{Osmolality}_{\text{Glucose}} = \frac{25}{180} = 0.139 \text{ Osm/kg} = 139 \text{ mOsm/kg}$$

First calculate mmolar sodium chloride concentration:

NaCl = 0.9% = 0.9 g/100 mL = 9 g/L

Final concentration (after mixing with an equal volume of 5% glucose) is one half of this i.e. 4.5 g/L. MW of NaCl = 23 + 35.5 = 58.5.

$$\text{NaCl (mol/L)} = \frac{\text{NaCL (g/L)}}{\text{MW}} = \frac{4.5}{58.5} = 0.077 \text{ mol/L} = 77 \text{ mmol/L}$$

Sodium chloride dissociates to give two osmotically active species – Na^+ and Cl^-

Therefore, $\text{Osmolality}_{\text{NaCl}}$ = 2 x 77 = 154 mOsm/kg

$$\text{Osmolality}_{\text{Total}} = \text{Osmolality}_{\text{Glucose}} + \text{Osmolality}_{\text{NaCl}}$$

$$= 139 + 154$$

$$= \mathbf{293\ mOsm/kg} \quad \text{(i.e. essentially isosmolar)}$$

2. *Calculate the approximate osmolality of a solution containing 10% mannitol and 0.9% saline (MW mannitol = 182).*

Calculate individual osmolalities separately.

Mannitol is undissociated so that $\text{Osmolality}_{\text{Mannitol}}$ = Molar concentration

Concentration of mannitol = 10% = 10 g/100 mL = 100 g/L

$$\text{Osmolality}_{\text{Mannitol}} = \frac{\text{Mannitol (g/L)}}{\text{MW}} = \frac{100}{182}\ \text{Osm/kg}$$

$$= \frac{100 \times 1{,}000}{182}\ \text{mOsm/kg} = 549\ \text{mOsm/kg}$$

Concentration of NaCl = 0.9% = 0.9 g/100 mL = 9 g/L

MW NaCl = 23 + 35.5 = 58.5

$$\text{NaCl (mmol/L)} = \frac{\text{NaCl (g/L)} \times 1{,}000}{\text{MW}} = \frac{9 \times 1{,}000}{58.5} = 154\ \text{mmol/L}$$

Each molecule of NaCl dissociates into 2 ions (Na^+ and Cl^-).

$\text{Osmolality}_{\text{NaCl}}$ = 2 x NaCl (mmol/L) = 2 x 154 = 308 mOsm/Kg

Adding these together gives the total osmolality:

$$\text{Osmolality}_{\text{Total}} = \text{Osmolality}_{\text{Mannitol}} + \text{Osmolality}_{\text{NaCl}}$$

$$= 549 + 308$$

$$= \mathbf{857\ mOsm/kg}$$

3. *A patient was mistakenly given 500 mL 20% mannitol ($C_6H_{14}O_6$) intended for the patient in the next bed instead of the same volume of normal (0.9%) saline. Calculate the extra osmolal load given over that which would have resulted from isotonic saline.*

First calculate the osmotic load of 500 mL of each solution.

For 20% mannitol:

Concentration = 20% = 20 g/ 100 mL = 200 g/L

AW C = 12, therefore C_6 = 6 x 12 = 72
AW H = 1, therefore H_{14} = 14 x 1 = 14
AW O = 16, therefore O_6 = 6 x 16 = 96
MW = 182

$$\text{Osmolality}_{\text{Mannitol}} = \frac{\text{Mannitol (g/L)}}{\text{MW}} = \frac{200}{182} \text{ Osm/kg}$$

$$= \frac{200 \times 1{,}000}{182} \text{ mOsm/kg} = 1099 \text{ mOsm/kg}$$

Therefore osmotic load of 500 mL = $\frac{1099}{2}$ = 550 mOsm

For 0.9% saline:

Concentration = 0.9% = 0.9 g/100 mL = 9 g/L

MW NaCl = 23 + 35.5 = 58.5

$$\text{Osmolality}_{\text{NaCl}} = \frac{\text{NaCl (g/L)} \times 2}{\text{MW}} \text{ Osm/kg} = \frac{9 \times 2}{58.5} \text{ Osm/kg}$$

$$= \frac{9 \times 2 \times 1{,}000}{58.5} \text{ mOsm/kg} = 308 \text{ mOsm/kg}$$

(factor of 2 introduced since NaCl dissociates into 2 ions (Na^+ and Cl^-).

Osmolar load due to 500 mL NaCl $= \frac{308}{2} = 154$ mOsm

Extra osmolal load = Osmolal load$_{Mannitol}$ - Osmol load$_{NaCl}$

= 550 - 154

= **396 mOsm**

4. *What increase in plasma osmoality would result from a plasma ethanol concentration of 92 mg/dL?*

First convert ethanol concentration to mmol/L:

Ethanol concentration = 92 mg/dL = 920 mg/L

Formula for ethanol = C_2H_5OH

AW C = 12, therefore C_2 = 2 x 12 = 24
AW H = 1, therefore H_6 = 6 x 1 = 6
AW O = 16, therefore O = 1 x 16 = 16
MW = 46

Osmolality$_{Ethanol}$ $= \frac{\text{Ethanol (mg/L)}}{\text{MW}} = \frac{920}{46} =$ **20 mOsm/kg**

5. *A 45-year old man is brought to casualty following a fit. He had been working alone late in a garage, when he was found by the security guard who called an ambulance. On admission, he has a large bruise on the left temple and is semi-comatose, he smells of alcohol. The admitting team request urea and electrolytes, glucose and an alcohol and blood gas estimation and arrange an urgent CT scan. The results are as follows:*

Sodium	*141 mmol/L*	*Urea*	*3.5 mmol/l*
Ethanol	*270 mg/dL*	*Glucose*	*3.2 mmol/L*

The CT scan does not show any bony injury or evidence of intracranial bleed. The neurological registrar is called and asks for an osmolal gap to help provide a quick estimation of whether there is a possibility that other toxic substances present in the garage, such as antifreeze, have been taken in any quantity.

The measured osmolality is 330 mOsm/kg

Calculate the osmolal gap
Show whether the alcohol concentration explains the observed osmolal gap, explaining any assumptions you make in the process.

a) First calculate osmolality due to Na^+, glucose and urea:

$$\underset{\text{mOsm/kg}}{\text{Osmolality}} = 1.86\,\underset{\text{mmol/L}}{[Na^+]} + \underset{\text{mmol/L}}{[\text{glucose}]} + \underset{\text{mmol/L}}{[\text{urea}]} + 9$$

$$\text{Osmolality} = (1.86 \times 141) + 3.2 + 3.5 + 9 = 278 \text{ mOsm/kg}$$

$$\text{Osmolal gap} = \text{Osmolality}_{\text{Measured}} - \text{Osmolality}_{\text{Calculated}}$$

$$= 330 - 278$$

$$= \mathbf{52\ mOsm/kg}$$

b) Calculate the expected contribution from ethanol:

Ethanol concentration = 270 mg/dL = 2,700 mg/L

Formula of ethanol = C_2H_5OH

AW C = 12, therefore C_2 = 2 x 12 = 24
AW H = 1, therefore H_6 = 6 x 1 = 6
AW O = 16, therefore O = 1 x 16 = 16
MW = 46

$$\text{Osmolality}_{\text{Ethanol}} = \frac{\text{Ethanol (mg/L)}}{\text{MW}} = \frac{2{,}700}{46} = 59 \text{ mOsm/kg}$$

The osmolal gap is in reasonable agreement with the expected osmolal contribution from ethanol. Therefore the ethanol concentration explains the observed osmolal gap.

Chapter 7

1. *An antidepressant drug has a biological half-life of 30 hours. How long will it take a plasma concentration of 50 mg/L to fall to 20 mg/L?*

The first order elimination rate equation is:

$$\ln Cp_t = \ln Cp_0 - k_d.t$$

Where Cp_t = drug concentration at time t = 20 mg/L
Cp_0 = initial drug concentration = 50 mg/L
k_d = elimination rate constant

k_d can be calculated from the half-life ($t_{½}$ = 30 h):

$$k_d = \frac{0.693}{t_{½}} = \frac{0.693}{30} = 0.023\ h^{-1}$$

Substitute these values into the rate equation and solve for t:

$$\ln 20 = \ln 50 - 0.023.t$$

$$3.00 = 3.91 - 0.023.t$$

$$0.023.t = 3.91 - 3.00 = 0.91$$

$$t = \frac{0.91}{0.023} = \textbf{40 h}\ (2\ \text{sig figs})$$

2. *A 15 year old boy presents to casualty following a convulsion. It turns out that he had swallowed 30 of his mother's lithium tablets about 10 hours previously. On admission his lithium concentration is 4.1 mmol/L. A decision needs to be made whether to haemodialyse him to reduce the lithium concentration. As this is not going to be available quickly, the physicians want to know how long he will have toxic levels just with endogenous clearance. Estimate the following, indicating clearly any assumptions you have made:*

   a) The likely volume of distribution of the lithium at this stage in the situation, given a body weight of 65 kg.

   b) How long it will be before his lithium concentration drops to the relatively safe level of 1.5 mmol/L below which toxicity is unlikely, given a clearance of 0.03 L/h/kg.

a) The volume of distribution of a drug is usually calculated by dividing the total dose administered by the plasma concentration. In this question we do not have a reliable estimate of the amount ingested. Since lithium is readily water soluble its volume of distribution approximates to total body water volume.

$$\text{Total body water (L)} = \text{Body wt (kg)} \times \frac{\%\ \text{Body water}}{100}$$

Assuming an average body water content of 60%:

$$\text{Volume of distribution } (V_d) = \frac{65 \times 60}{100} = \mathbf{39\ L}$$

c) Lithium is excreted from the body by glomerular filtration (with some reabsorption by the proximal tubule which we shall ignore) and so its elimination follows first order kinetics:

$$\ln Cp_t = \ln Cp_0 - k_d.t$$

Cp_0 = initial concentration (before dialysis) = 4.1 mmol/L

Cp_t = concentration at time t = 1.5 mmol/L

t = time taken (in hours) to reach the "safe" level of 1.5 mmol/L

k_d = elimination rate constant

The clearance of the drug is given as 0.03 L/h/kg. Multiply by the patient's weight to give the total clearance:

Clearance = 0.03 (L/h/kg) x 65 (kg) = 1.95 L/h

The elimination rate constant (k_d) can be calculated from the clearance (Cl) and the volume of distribution (V_d):

$$k_d = \frac{Cl}{V_d} = \frac{1.95}{39} = 0.050\ h^{-1}$$

Substitute for Cp_t, Cp_0 and k_d then solve for t:

$$\ln 1.5 = \ln 4.1 - 0.050.t$$

$$0.405 = 1.411 - 0.050.t$$

$$0.050.t = 1.411 - 0.405 = 1.006$$

$$t = \frac{1.006}{0.050} = \mathbf{20\ h} \text{ (2 sig figs)}$$

3. *A 60 mg dose of a drug is given to a male experimental subject who weighs 80 kg. Assuming that the drug is completely absorbed and distributed evenly throughout the total body water, estimate the potential peak plasma level. If the drug were distributed only within the extracellular compartment, what would the plasma level be?*

Assuming distribution throughout total body water, the V_d = total body water vol:

Assume body water is 60% of body weight.

Total body water (L) = Body Wt (Kg) x 60%

$$= \frac{80 \times 60}{100} = 48 \text{ L}$$

$$V_d = \frac{\text{Amount of drug in body (dose)}}{\text{Plasma drug concentration}}$$

$$\text{Plasma drug level (mg/L)} = \frac{\text{Dose (mg)}}{V_d \text{ (L)}} = \frac{60}{48} = \mathbf{1.25 \text{ mg/L}}$$

If drug is only distributed throughout the ECF, the V_d must be adjusted. ECF is normally 20% of body wt.

V_d (L) = Body wt (kg) x 20%

$$= \frac{80 \times 20}{100} = 16 \text{ L}$$

$$\text{Plasma drug level (mg/L)} = \frac{\text{Dose (mg)}}{V_d \text{ (L)}} = \frac{60}{16} = \mathbf{3.75 \text{ mg/L}}$$

Alternatively, since a third of body water is in the ECF, the drug level will be 3 times higher.

4. A bolus of 6 g drug is given IV and 3 blood samples collected at intervals.

Time	*mg/L*
2.5h	*32*
5h	*10*
7.5h	*3*

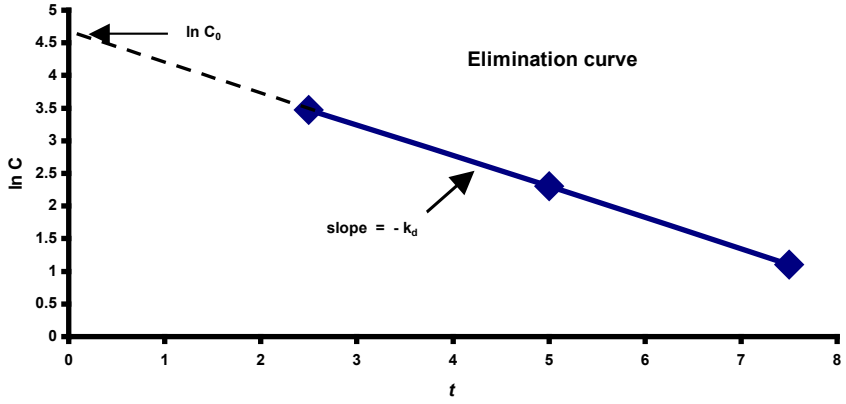
a) What is the half-life of the drug?
b) What is the volume of its distribution?

a) Assuming the clearance of the drug follows first order elimination kinetics then the data should be described by the expression:

$$\ln Cp_t = \ln Cp_0 - k_d.t$$

Therefore a plot of ln *C* versus *t* should be linear with an intercept on the ln *C* axis of Cp_0 and slope $-k_d$:

Time (h)	Conc (mg/L)	ln*C*
2.5	32	3.47
5	10	2.30
7 .5	3	1.10



This plot clearly demonstrates that elimination of the drug follows first order kinetics so that Cp_0 and k_d could be determined directly from the graph. Alternatively any 2 values can be substituted into the rate equation and solved for k_d:

Let level at time 2.5 h be Cp_0 = 32 mg/L
Let level at time 5 h be Cp_t = 10 mg/L
t = 2.5 h (the time difference between Cp_0 and Cp_t).

Therefore: $\ln 10 = \ln 32 - k_d.2.5$

$$2.303 = 3.466 - 2.5\ k_d$$

$$2.5\ k_d = 3.466 - 2.303 = 1.163$$

$$k_d = \frac{1.163}{2.5} = 0.465\ h^{-1}$$

a) Half-life ($t_{½}$) can be calculated from k_d:

$$t_{½} = \frac{0.693}{k_d} = \frac{0.693}{0.465} = \mathbf{1.5\ h} \quad (2\ sig\ figs)$$

b) First calculate the initial concentration (Cp_0) using one other value (e.g. 2.5 h = 32 mg/L as Cp_t and t = 2.5 h) and the value for k_d:

$$\ln 32 = \ln Cp_0 - (0.465 \times 2.5)$$

$$3.466 = \ln Cp_0 - 1.163$$

$$\ln Cp_0 = 3.466 + 1.163 = 4.629$$

$$Cp_0 = \text{antilog}_e\ 4.629 = 102\ mg/L$$

The V_d is then calculated from dose and Cp_0:

$$V_d\ (L) = \frac{\text{Dose (mg)}}{Cp_0\ (mg/L)} = \frac{6{,}000}{102} = \mathbf{59\ L} \quad (2\ sig\ figs)$$

5. *The plasma concentration of a drug is found to be 200 nmol/L at 9.00 am. It's elimination follows first order kinetics with a rate constant is 0.34/h. Calculate the times at which the plasma concentrations will reach 100 nmol/L and 75 nmol/L.*

The first order rate equation is:

$$\ln Cp_t = \ln Cp_0 - k_d.t$$

where Cp_0 = 200 nmol/L; k_d = 0.34 h^{-1}

Calculation of t when Cp_t = 100 nmol/L:

$$\ln 100 = \ln 200 - 0.34.t$$

$$4.605 = 5.298 - 0.34.t$$

$$0.34.t = 5.298 - 4.605 = 0.693$$

$$t = \frac{0.693}{0.34} = \textbf{2.0 h} \quad (2 \text{ sig figs})$$

Calculation of t when Cp_t = 75 nmol/L

$$\ln 75 = \ln 200 - 0.34.t$$

$$4.317 = 5.298 - 0.34.t$$

$$0.34.t = 5.298 - 4.317 = 0.981$$

$$t = \frac{0.981}{0.34} = \textbf{2.9 h} \quad (2 \text{ sig figs})$$

6. *A patient in casualty with a suspected adrenal crisis is given an iv dose of hydrocortisone at 18.00. The medical team on take wish to carry out a short synacthen test to confirm the diagnosis but there will be a significant contribution form the administered drug until its concentration has fallen to 10% of the peak value. If the half-life of hydrocortisone is 2 h, what is the earliest time at which the test can be carried out?*

Assuming elimination follows first order kinetics:

$$\ln Cp_t = Cp_0 - k_d.t$$

Where Cp_t = concentration at time t = 10%
Cp_0 = initial concentration = 100%
t = time when Cp_t reaches 10%

Calculate k_d from $t_{½}$:

$$k_d = \frac{0.693}{t_{½}} = \frac{0.693}{2} = 0.347\ h^{-1}$$

Substitute these values into the rate equation and solve for t:

$$\ln 10 = \ln 100 - 0.347.t$$

$$2.303 = 4.605 - 0.347.t$$

$$0.347.t = 4.605 - 2.303 = 2.302$$

$$t = \frac{2.302}{0.347} = \mathbf{6.6\ h}\ (2\ \text{sig figs})$$

7. *The SHO decides to treat a patient (weight 80 kg) with intravenous theophylline (salt factor = 0.8). What loading dose would you recommend in order to achieve a theophylline level of 12 mg/L given a volume of distribution of 0.5 L/kg and an initial plasma theophylline level of 4 mg/L?*

$$\text{Loading dose } (LD) = \frac{V_d \times (Cp_{target} - Cp_{initial})}{S \times F}$$

Where V_d = volume of distribution = 0.5 L/kg
Cp_{target} = desired drug level = 12 mg/L
$Cp_{initial}$ = starting drug level = 4 mg/L
S = salt factor = 0.8
F = bioavailability (not given so assume a value of 1)

Patients V_d = Body weight (kg) x 0.5

= 80 x 0.5 = 40 L

Substitute these values in order to obtain *LD*:

$$LD = \frac{40 \times (12 - 4)}{0.8}$$

$$= \frac{40 \times 8}{0.8}$$

= **400 mg**

8. *A patient, (body weight 55 kg) unable to take oral medication, had been receiving intravenous valproate for a number of days and achieved an average steady state level of 75 mg/L. After regaining consciousness the doctors wished to change to an oral twice daily regimen. In order to maintain the same average steady state concentration what dose would you recommend. Assume a clearance of 10 mL/h/kg, a bioavailability of 0.7 and a salt factor of 0.85.*

The following expression allows calculation of the maintenance dose:

$$\text{Maintenance dose} = \frac{Cp_{ss} \times Cl \times \tau}{S \times F}$$

Where:

Cp_{ss} = steady state plasma concentration = 75 mg/L
Cl = clearance = 10 mL/h/kg
τ = dosing interval = 12 h (i.e. twice daily)
S = salt factor = 0.85
F = bioavailability = 0.7

First correct the clearance for the body weight and express it in litres (to be compatible with the drug concentration which is given in mg/L):

$$Cl \text{ (L/h)} = \frac{Cl \text{ (mL/h/kg)} \times \text{Body wt (kg)}}{1{,}000}$$

$$= \frac{10 \times 55}{1{,}000} = 0.55 \text{ L/h}$$

Substitute these values into the expression for maintenance dose:

$$\text{Maintenance dose (mg)} = \frac{75 \times 0.55 \times 12}{0.85 \times 0.7}$$

$$= \mathbf{832 \text{ mg}}$$

Chapter 8

1. *Over a 24 h period a patient recovering from intestinal resection receives 2 L of fluids intravenously and 750 mL orally but does not eat any solids over this period. The urine output over the same period is 1.25 L and 600 mL of fluid is lost via a fistula. Is the patient in positive or negative fluid balance and by how much?*

Draw up a table of fluid gains and losses then calculate the total of each. Assume a value of 400 mL per day for *net insensible losses*.

	Fluid gains		**Fluid losses**	
	Oral	750 mL	Urine output	1,250 mL
	IV	2,000 mL	Loss via fistula	600 mL
			Net insensible loss	400 mL
Total		2,750 mL		2,250 mL

Fluid balance (mL) = Net fluid intake (mL) - Net fluid loss (mL)

= 2,750 - 2,250

= **500 mL**

i.e. there is a net fluid gain of 500 mL.

2. *A patient known to have diabetes insipidus is admitted in coma. His plasma osmolality is 324 mosm/kg. If his weight is 85 kg, estimate his body water deficit.*

The average adult male has a body water content of approximately 60%. If the body water deficit is x L, then the initial body water content can be calculated:

$$\text{Initial body water (L)} = \frac{(85 + x) \times 60}{100}$$

$$= \frac{5100 + 60x}{100}$$

$$= 51 + 0.6x$$

Assuming a normal initial osmolality (say 285 mOsm/kg) the total amount (in mOsm) of osmotically active species present in the body can be calculated:

$$\text{Osmolality (mOsm/kg)} = \frac{\text{Total solutes (mOsm)}}{\text{Initial body water (kg)}}$$

$$285 = \frac{\text{Total solutes (mOsm)}}{51 + 0.6x}$$

$$\text{Total solutes (mOsm)} = 285\,(51 + 0.6x)$$

$$= 14{,}535 + 171x$$

On presentation his body weight is 85 kg. Assuming the total amount of solutes in the body is unchanged, then the body water volume can be calculated from the current osmolality:

$$\text{Final osmolality (mOsm/kg)} = \frac{\text{Total solutes (mOsm)}}{\text{Final body water (kg)}}$$

$$324 = \frac{14{,}535 + 171x}{(51 + 0.6x) - x}$$

$$324 = \frac{14{,}535 + 171x}{51 - 0.4x}$$

$$324(51 - 0.4x) = 14{,}535 + 171x$$

$$16{,}524 - 130x = 14{,}535 + 171x$$

$$171x + 130x = 16{,}524 - 14{,}535$$

$$301x = 1989$$

$$x = \frac{1989}{301}$$

$$= \mathbf{6.6\ L}$$

If it is assumed that the change in body wt is neglible (or that the initial body water was the same as for an average 70 kg adult) then a simpler calculation (using Eq. 8.3) can be used and gives a slightly different result which may be adequate as a rough guide in clinical practice:

$$\text{Fluid loss (L)} = 42 - \left[\frac{12000}{\text{Osmolality (mOsm/kg)}}\right]$$

$$= 42 - \frac{[12000]}{324}$$

$$= 42 - 37$$

$$= \mathbf{5L}$$

3 *A male adult insulin dependent diabetic forgot to take his insulin. His blood glucose concentration, which was 5 mmol/L, rose to 15 mmol/\|L in two hours. Estimate the effect on his plasma sodium concentration, assuming that no other water intake nor loss of water from the body takes place during this time, indicating what assumptions you make.*

Making a number of assumptions:

- That it is plasma glucose which is measured rather than whole blood glucose.
- That as a result of insulin deficiency there is no increase in glucose concentration in the intracellular fluid (ICF).
- That the plasma glucose has equilibrated with interstitial fluid so that it's concentration in the extracellular fluid (ECF) is the same as in plasma.
- That there is negligible change in the concentrations of solutes other than glucose, sodium and chloride.
- That the ratio of ICF:ECF volumes is 2 (i.e. ECF = 14L, ICF = 28L for average adult male) and that the total body water is that of an average male i.e. 42 L

The effect of an increase in plasma (and hence ECF) glucose is to raise plasma (and ECF) osmolarity. The body will retain water (stimulation of thirst increases intake and stimulation of ADH reduces renal loss) until osmotic equilibrium is restored. If there is a plentiful supply of water then the plasma osmolarity is returned to normal and since the plasma glucose has risen by 10 mmol/L the plasma sodium must have fallen by 10/2 = 5 mmol/L. However, this question states that **there is no net loss or gain of body water**. Therefore, water will move, by osmosis, from the ICF compartment (iso-osmolar) to the ECF (now hyper-osmolar) until osmotic equilibrium is established. Since movement of water from the ICF leads to an increase in ICF osmolarity, the movement of water is restricted and at equilibrium the ECF will reach a value somewhere in-between normality and the original value i.e. the osmotic load is shared between the ECF and ICF compartments, both of which become hyperosmolar.

The plasma glucose has risen by 15 - 5 = 10 mmol/L

Rise in **amount** of glucose in ECF =

Rise in plasma glucose concentration (mmol/L) x ECF vol(L)

= 10 x 14 = 140 mmol

(a slight underestimate since there has been a small expansion in ECF vol)

At equilibrium, the rise in osmolarity (which is the same in the ECF and ICF) is given by:

$$\frac{\text{Increase in amount of glucose in body (mmol)}}{\text{Total body fluid (ECF + ICF) volume (L)}}$$

$$= \frac{140}{42} = 3.33 \text{ mmol/L}$$

Since the plasma osmolarity has risen by 3.33 mOsmol/L and the plasma glucose by 10 mmol/L then the concentration of NaCl which has been displaced by glucose is

10 - 3.33 = 6.67 mmol/L

and so the sodium has fallen by $\frac{6.67}{2}$ = 3.34 mmol/L

i.e. the plasma sodium concentration has decreased by approximately **3 mmol/L**.

4. *A plasma sample with a total protein content of 70 g/L gave identical sodium results of 140 mmol/L when measured using either a direct-reading ion-selective electrode or a flame photometer. What plasma sodium result would you expect the ion-selective electrode to give with the same plasma sample if its total protein concentration had been 90 g/L?*

Flame photometry measures sodium as concentration in plasma i.e. 140 mmol/L of **plasma**.

A direct-reading ion-selective electrode measures sodium as activity i.e. 140 mmol/L of plasma **water**. Large molecules such as proteins occupy significant space in solution i.e. displace plasma water. If plasma contains 70 g/L protein then this is equivalent to 0.070 kg/L. Assuming that 1 kg of protein occupies a volume of 1 L then the volume of plasma water in which the sodium is dissolved is (1.0 - 0.07) = 0.93 L. Assuming that the activity is the same as concentration for sodium in plasma water (i.e. the activity coefficient is one), for a plasma sodium of 140 mmol/L of plasma, the true concentration of sodium in plasma water is:

$$\text{Plasma sodium} = \frac{140}{0.93} = 150.5 \text{ mmol/L water}$$

There are two ways in which the ISE reading can be converted to the same as that obtained by flame photometry (140 mmol/L):

- Subtraction of 10.5 mmol/L from the result
- Multiplication of the result by the factor 140/150.5 i.e. 0.930

At a protein concentration of 90 g/L (occupying 0.090 L plasma), the concentration of sodium in plasma water will be:

$$\frac{140}{(1.00 - 0.09)} = \frac{140}{0.91} = 153.8 \text{ mmol/L plasma water}$$

Carrying out the two adjustments by the instrument:

- Subtraction of 10.5 gives 153.8 - 10.5 = 143.3 mmol/L.
- Multiplication by 0.930 gives 153.8 x 0.930 = 143.0 mmol/L

Therefore expected ISE reading = **143 mmol/L**

Chapter 9

1. *An assay mixture for the measurement of lactate dehydrogenase constituted 2.7 mL of buffered NADH and 100 µL of serum. The reaction was started by adding 100 µL of sodium pyruvate. The absorbance change over 5 minutes was 0.150 when measured in a 0.5 cm light path at 340 nm. Assuming the molar absorbtivity of NADH at 340 nm is 6.30 x 10^3 $L.mol^{-1}cm^{-1}$, calculate the enzyme activity in international units per litre of serum.*

One international unit of activity is the amount of enzyme present in 1 L of serum which catalyses the conversion of 1 µmol substrate per min under the conditions of the assay.

First calculate the absorbance change per min:

$$\Delta A/\text{min} = \frac{\Delta A/5\text{min}}{5} = \frac{0.150}{5}\ \Delta A/\text{min}$$

Convert the absorbance change to concentration change per min:

$$\Delta A = a.b.\Delta c$$

Where $\Delta A/\text{min} = \frac{0.150}{5}$

a = molar absorptivity = 6.30 x 10^3 $L.mol^{-1}.cm^{-1}$

b = light path length = 0.5 cm

Δc = change in concentration (mol/min)

Substitute these values then re-arrange to give an expression for Δc/min:

$$\frac{0.150}{5} = 6.30 \times 10^3 \times 0.5 \times \Delta c/\text{min}$$

$$\Delta c/\text{min} = \frac{0.150}{5 \times 6.30 \times 10^3 \times 0.5}\ \text{mol/min/L reaction mixture}$$

Multiply by 1,000,000 to convert the concentration units from mol to µmol (1 mol = 1,000,000 µmol):

$$\Delta c/\text{min} = \frac{0.150 \times 1{,}000{,}000}{5 \times 6.30 \times 10^3 \times 0.5} \text{ µmol/min/L reaction mixture}$$

The final step is to convert the activity to µmol/min/L serum. In the assay 100 µL of serum was mixed with 2.7 mL buffer and 100 µl of substrate.

$$\text{Total assay volume} = 2.7 + 0.1 + 0.1 = 2.9 \text{ mL}$$

$$\Delta c/\text{min/L serum} = \frac{\Delta c/\text{min/L assay mixture} \times \text{Total assay vol (mL)}}{\text{Serum vol (mL)}}$$

$$= \frac{0.150 \times 1{,}000{,}000 \times 2.9}{5 \times 6.30 \times 10^3 \times 0.5 \times 0.1}$$

$$= \mathbf{276 \text{ IU/L serum}}$$

2. *An assay for alkaline phosphatase activity involved mixing 0.05 mL of serum with 2.7 mL buffer, allowing temperature to reach equilibrium then starting the reaction by adding 0.2 mL of substrate (4-nitrophenyl phosphate). The increase in absorbance in a 1cm cuvette due to the liberation of product (4-nitrophenol) was 0.180 over a 5-minute period. Calculate the alkaline phosphtase activity expressing the result as a) international units per litre of serum, and b) katals per litre of serum. Assume that the molar absorptivity of 4-nitrophenol is 1.88 x 10^4 L/mol/cm.*

a) One international, unit is the amount of enzyme which liberates one µmol of product per minute. Therefore to calculate the alk phos activity in IU/L serum the following steps are involved:

Determine the rate of absorbance change in ΔA/min.

$$\Delta A/\text{min} = \frac{\Delta A/5\text{min}}{5} = \frac{0.180}{5}$$

Convert to the rate of change in concentration using the molar absorptivity and pathlength:

$$\Delta A = a.b.\Delta c$$

ΔA = rate of absorbance change = $\frac{0.180}{5}$ A/min

a = molar absorptivity of 4-nitrophenol = 1.88×10^4 L/mol/cm

b = pathlength of cuvette = 1 cm

Δc = rate of change of concentration in mol/L/min = ?

Substitute these values and re-arrange to give an expression for Δc/min:

$$\frac{0.180}{5} = 1.88 \times 10^4 \times 1 \times \Delta c/\text{min}$$

$$\Delta c/\text{min} = \frac{0.180}{5 \times 1.88 \times 10^4 \times 1} \text{ mol/L/min}$$

Multiply by 1,000,000 to convert units from mol/L/min to µmol/L/min (1 mol = 1,000,000 µmol):

$$\Delta c/\text{min} = \frac{0.180 \times 1{,}000{,}000}{5 \times 1.88 \times 10^4 \times 1} \text{ µmol/min/L reaction mixture}$$

To convert to activity per L serum multiply by the total volume of reaction mixture and divide by the sample volume – using the same units:

Serum	=	0.05 mL
Buffer	=	2.70 mL
Substrate	=	0.20 mL
Total	=	2.95 mL

$$\text{Alk phos activity} = \frac{0.180 \times 1{,}000{,}000 \times 2.95}{5 \times 1.88 \times 10^4 \times 1 \times 0.05} \text{ µmol/min/L serum}$$

$$= \mathbf{113\ IU/L}$$

b) One Katal is the amount of enzyme that catalyses the reaction of 1 mol substrate per second

If alk phos activity = 113 IU/L (µmol/min/L)

Then divide by 1,000,000 to convert from µmol to mol, then by 60 to convert from min to seconds.

$$113 \text{ IU/L} = \frac{113}{1{,}000{,}000 \times 60} \text{ Katals/L}$$

$$= \mathbf{1.88 \times 10^{-6} \text{ Kat/L}}$$

3. *The Somogyi saccharogenic method for the assay of amylase involves measuring the rate of release of glucose from substrate. One Somogyi unit is the amount of enzyme catalysing the release of 1 mg of glucose in 30 min per 100 mL serum. Derive a factor to convert Somogyi units to international units per litre of serum.*

Somogyi units = mg glucose/30 min/100 mL serum

International units = µmol/min/L serum

Consider a sample with activity of *x* Somogyi units

First convert from mg glucose to µmol glucose:

Glucose formula = $C_6H_{12}O_6$

AW C = 12, therefore C_6 = 6 x 12 = 72
AW H = 1, therefore H_{12} = 12 x 1 = 12
AW O = 16, therefore O_6 = 6 x 16 = 96
MW = 180

Activity (Somogyi units) = *x* mg/30 min/100 mL

$$\text{Activity (mmol/30/min/100 mL)} = \frac{x}{180}$$

Multiply by 1,000 to convert from mmol to µmol:

$$\text{Activity (µmol/30 min/100 mL)} = \frac{x \times 1{,}000}{180}$$

Divide by 30 to obtain the rate per minute:

$$\text{Activity (µmol/min/100 mL)} = \frac{x \times 1{,}000}{180 \times 30}$$

Multiply by 10 to obtain the activity per litre:

$$\text{Activity (µmol/min/L)} = \frac{x \times 1{,}000 \times 10}{180 \times 30}$$

$$= x \times 1.85$$

$$\text{IU/L} = \mathbf{1.85} \times \text{Somogyi Units}$$

4 *One Wroblewski-laDue unit is the amount of lactate dehydrogenase which results in an absorbance change (due to NADH) at 340 nm of 0.001 per minute per mL serum in a reaction mixture with a total volume of 3 mL. Derive a factor to convert Wroblewski-LaDue units to International units per litre of serum. Assume the molar absorptivity of NADH is 6.3 x 10^3 L/mol/cm.*

International units = µmol/min/L serum

Wroblewski-laDue (W-l-D units) = 0.001 ΔA/min/mL serum (total volume 3 mL)

Multiply by 3 to obtain absorbance change obtained with 1 mL of undiluted serum, then by 1,000 to obtain the absorbance change due to 1 L serum:

1 W-l-D unit = 0.001 x 3 x 1,000 ΔA/min/L serum

Next convert ΔA/min to Δc (i.e. mol/L/min):

$$\Delta A = a.b.\Delta c$$

ΔA = absorbance change = 0.001 x 3 x 1,000 A/min

a = molar absorptivity of NADH = 6.3×10^3 $L.mol^{-1}.cm^{-1}$

b = cuvette pathlength = not given so assume 1 cm

Δc = rate of change of concentration in mol/L/min

Substitute these values and re-arrange to give an expression for Δc:

$$0.001 \times 3 \times 1{,}000\ A/\text{min} = 6.3 \times 10^3 \times 1 \times \Delta c$$

$$\Delta c = \frac{\Delta A \times 3 \times 1{,}000}{6.3 \times 10^3 \times 1} \text{ mol/min/L serum}$$

Multiply by 1,000,000 to convert from mol to µmol:

$$1 \text{ W-l-D unit} = \frac{0.001 \times 1{,}000 \times 3 \times 1{,}000{,}000}{6.3 \times 10^3 \times 1} \text{ µmol/min/L reaction mixture}$$

1 W-l-D unit = 476 µmol/min/L serum

i.e. 1 W-l-D unit = 476 IU/L

Therefore activity (IU/L) = **476** x Wroblewski-laDue Units

5. *If the K_m of an enzyme which obeys simple Michaelis-Menten kinetics is 2.5 mmol/L, what velocity (expressed as a multiple of V_{max}) would be obtained at a substrate concentration of 10 mmol/L?*

The Michaelis –Menten equation is:

$$v = \frac{V_{max}[S]}{K_m + [S]}$$

Where v = initial velocity
V_{max} = maximal velocity
[S] = substrate concentration = 10 mmol/L
K_m = Michaelis constant = 2.5 mmol/L

Substitute and solve for v:

$$v = \frac{10\ V_{max}}{2.5 + 10} = \frac{10\ V_{max}}{12.5} = \mathbf{0.8\ V_{max}}$$

6. *What information can be obtained from the double-reciprocal plot for an enzyme under the following conditions: a) 1/v = 0 when 1/[S] = -12.5 x 10^6 L/mol, b) 1/[S] = 0 when 1/v = 5.2 x 10^6 min/mol, c) 1/[S] = 0 when 1/v = 6.5 x 10^6 min/mol and the slope of the line is 100 min/L?*

a) When 1/v = 0, the value for 1/[S] is $-1/K_m$

Therefore $\frac{-1}{K_m}$ = -12.5 x 10^6 L/mol

Which can be re-arranged to give the value of K_m:

$$\mathbf{K_m} = \frac{-1}{-12.5 \times 10^6} = 0.08 \times 101^{-6} = \mathbf{8.0 \times 10^{-8}\ mol/L}$$

b) When $1/[S] = 0$, $1/v = 1/V_{max}$

Therefore $\frac{1}{V_{max}} = 5.2 \times 10^6$ min/mol

Rearrange and solve for V_{max}:

$$\boldsymbol{V_{max}} = \frac{1}{5.2 \times 10^6} = 0.19 \times 10^{-6} = \mathbf{1.9 \times 10^{-7}\ mol/min}$$

c) When $1/[S] = 0$, $1/v = 1/V_{max}$

Therefore $\frac{1}{V_{max}} = 6.5 \times 10^6$ min/mol

Rearrange and solve for V_{max}:

$$V_{max} = \frac{1}{6.5 \times 10^6} = 0.15 \times 10^{-6} = \mathbf{1.5 \times 10^{-7}\ mol/min}$$

The slope of the line gives K_m/V_{max}

Therefore $\frac{K_m}{V_{max}} = 100$ min/L

Substitute $V_{max} = 1.5 \times 10^{-7}$ mol/min and solve for K_m:

$$\frac{K_m}{1.5 \times 10^{-7}} = 100$$

$$\boldsymbol{K_m} = 100 \times 1.5 \times 10^{-7} = \mathbf{1.5 \times 10^{-5}\ mol/L}$$

7. *You carry out an enzyme experiment in which the substrate concentration is expressed as mmol/L and the reaction velocity in µmol/L/min. What would be the units for the axes of the three following plots: a) 1/[S] versus 1/v, b) [S]/v versus [S], c) v versus v/[S]?*

a) If [S] = mmol/L = 10^{-3} mol/L

$$\frac{1}{[S]} = \frac{1}{10^{-3}\ \text{mol/L}} = \mathbf{10^3\ L/mol}$$

If v = µmol/min = 10^{-6} mol/min

$$\frac{1}{v} = \frac{1}{10^{-6}\ \text{mol/min}} = \mathbf{10^6\ min/mol}$$

b) $$\frac{[S]}{v} = \frac{\text{mmol/L}}{\text{µmol/min}} = \frac{10^{-3}\ \text{mol/L}}{10^{-6}\text{mol/min}} = \mathbf{10^3\ min/L}$$

[S] = mmol/L = **10^{-3} mol/L**

c) v = µmol/min = **10^{-6} mol/min**

$$\frac{v}{[S]} = \frac{\text{µmol/min}}{\text{mmol/L}} = \frac{10^{-6}\ \text{mol/min}}{10^{-3}\ \text{mol/L}} = \mathbf{10^{-3}\ L/min}$$

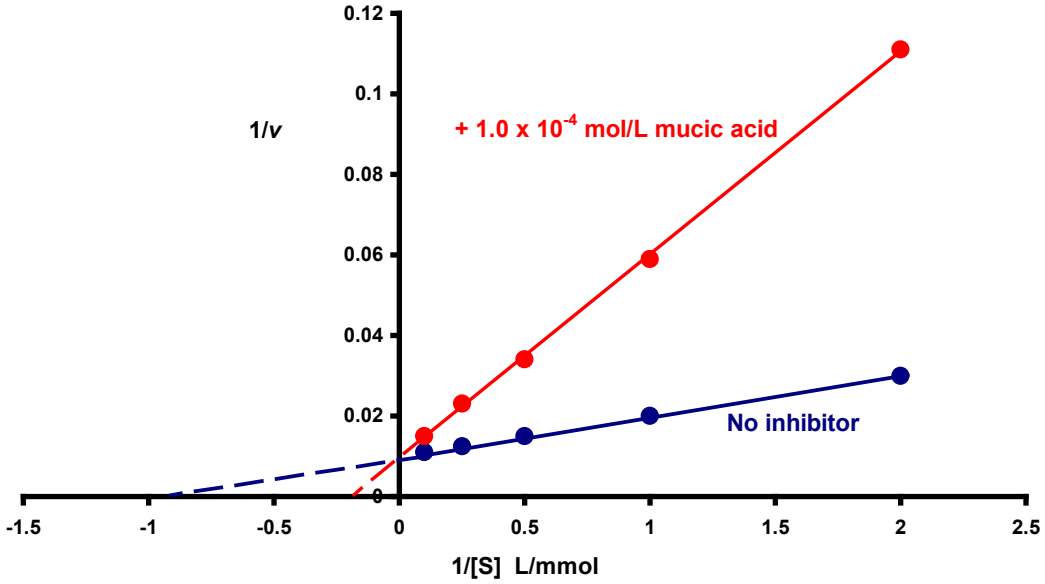
8. *Mucic acid is an inhibitor of β-glucuronidase. The following data were obtained using phenolphthalein glucuronide as substrate, in the presence and absence of mucic acid (concentration in the assay = 1.0 x 10^{-4} mol/L).*

Substrate Concentration (mmol/L)	*Reaction velocity*	
	No inhibitor	*Mucic acid*
0.5	*33*	*9*
1.0	*50*	*17*
2.0	*67*	*29*
4.0	*80*	*44*
10	*91*	*67*

Determine the type of inhibition and the enzyme-inhibitor dissociation constant.

The first step is to plot the data. Any linear transformation of the Michaelis-Meneten equation can be used but the double-reciprocal plot is probably the simplest. Calculated reciprocals are:

1/[S]	1/*v*	
L/mmol	No inhibitor	Mucic acid
2.0	0.030	0.111
1.0	0.020	0.059
0.5	0.015	0.034
0.25	0.0125	0.023
0.10	0.011	0.015



Without inhibitor, when $1/v = 0$, $1/[S] = -1/K_m$

Intercept on 1/[S] without inhibitor = - 0.947 L/mmol

Therefore $K_m = \dfrac{-1}{-0.947} = 1.06$ mmol/L $= 1.06 \times 10^{-3}$ mol/L

Since the lines cross on the $1/v$ axis the type of inhibition is **competitive**.

With inhibitor, when $1/v = 0$, $1/[S] = -1/K_m^{app}$

Intercept on 1/[S] with inhibitor = - 0.193 L/mmol

Therefore $K_m^{app} = \dfrac{-1}{-0.193} = 5.18$ mmol/L $=$ **5.18×10^{-3} mol/L**

For competitive inhibition: $K_m^{app} = K_m (1 + [I]/K_i)$

Substitute K_{mapp} = 5.18 x 10^{-3} mol/L, K_m = 1.06 x 10^{-3} mol/L, [I] = 1.0 x 10^{-4} mol/L then solve for K_i:

$$5.18 \times 10^{-3} = 1.06 \times 10^{-3}\{1 + (1.0 \times 10^{-4}/K_i)\}$$

$$\frac{5.18 \times 10^{-3}}{1.06 \times 10^{-3}} = 1 + \frac{1.0 \times 10^{-4}}{K_i}$$

$$4.89 - 1 = \frac{1.0 \times 10^{-4}}{K_i}$$

$$K_i = \frac{1.0 \times 10^{-4}}{3.89} = \mathbf{2.6 \times 10^{-5}\ mol/L}$$

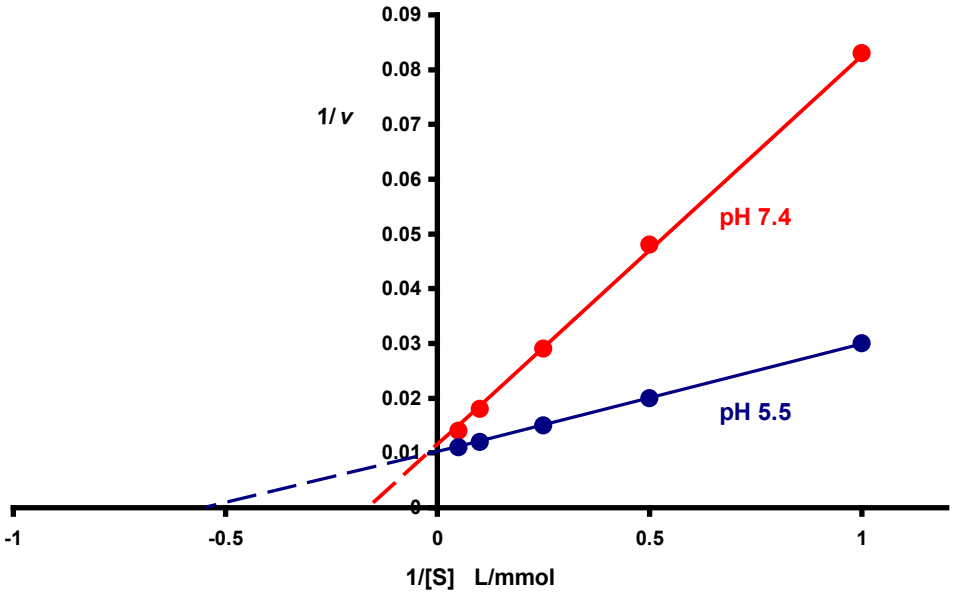
9. *An experiment was conducted to study the effect of pH on the activity of lactate dehydrogenase. Using a histidine buffer at pH 5.5 and 7.4 the reaction was monitored by following the increase in absorbance at 340 nm due to the reduction of NAD. The following data were obtained:*

Lactate concentration mmol/L	*Reaction velocity*	
	pH 7.4	*pH 5.5*
1	*12*	*33*
2	*21*	*50*
4	*35*	*67*
10	*57*	*83*
20	*73*	*91*

Stating any assumptions that you make determine the pH at which the enzyme has greatest affinity for the substrate.

Calculate reciprocals then plot 1/*v* versus 1/[S] at each pH:

1/[S] L/mmol	1/*v* pH 7.4	pH 5.5
1	0.083	0.030
0.5	0.048	0.020
0.25	0.029	0.015
0.1	0.018	0.012
0.05	0.014	0.011



When $1/v = 0$, $1/[S] = -1/K_m$

At pH 7.4, when $1/v = 0$, $1/[S] = -0.162$ L/mmol

Therefore $K_m = \frac{-1}{-0.162} = 6.2$ mmol/L = **6.2×10^{-3} mol/L**

At pH 5.5, when $1/v = 0$, $1/[S] = -0.545$ L/mmol

Therefore $K_m = \frac{-1}{-0.545} = 1.84$ mmol/L = **1.84×10^{-3} mol/L**

Assuming that equilibrium conditions apply (i.e. that $k_{+1} >> K_{+2}$) then the K_m is the dissociation constant of the enzyme-substrate complex and is inversely proportional to the affinity of the enzyme for the substrate. The Km is lower at pH 5.5 than at pH 7.4. **Therefore the enzyme has greatest affinity for its substrate at pH 5.5.**

10. *The apparent K_m and V_{max} of an enzyme were measured over a range of inhibitor concentrations and the following data obtained:*

Inhibitor concentration (mmol/L)	*Apparent value* K_m *(mmol/L)*	V_{max} *(µmol/min)*
5	*10*	*7.5*
10	*7*	*5*
15	*5*	*4*
20	*4*	*3*

Determine the mode of inhibition and the inhibitor constant (K_i).

A competitive inhibitor causes an increase in the apparent K_m. As the K_m is actually decreasing as inhibitor concentration increases this mode of inhibition can be ruled out. The apparent V_{max} is decreasing as inhibitor concentration is increased; this behaviour is seen both with non-competitive and uncompetitive inhibition. However, in non-competitive inhibition the K_m is unaffected by the inhibitor whereas in uncompetitive inhibition the apparent K_m decreases with increasing inhibitor concentration. Therefore these data are consistent with **uncompetitive inhibition**.

The value for K_i can be obtained from secondary plots of either $1/K_m$ or $1/V_{max}$ versus [I].

The relationship between K_m^{app} and [I] for an uncompetitive inhibitor is:

$$K_m^{app} = \frac{K_m}{(1 + [I]/K_i)}$$

Inversion of this expression gives:

$$\frac{1}{K_m^{app}} = \frac{(1 + [I]/K_i)}{K_m}$$

Which can also be written:

$$\frac{1}{K_m^{app}} = \left(\frac{1}{K_i K_m} \times [I] \right) + \frac{1}{K_m}$$

Therefore a plot of $1/K_m^{app}$ versus [I] is linear.

When $1/K_m^{app} = 0$:

$$0 = \left(\frac{1}{K_i K_m} \times [I] \right) + \frac{1}{K_m}$$

Which can be rearranged to give:

$$-\frac{1}{K_m} = \frac{[I]}{K_i K_m}$$

Multiplying both sides by $K_m K_i$ and changing the signs gives:

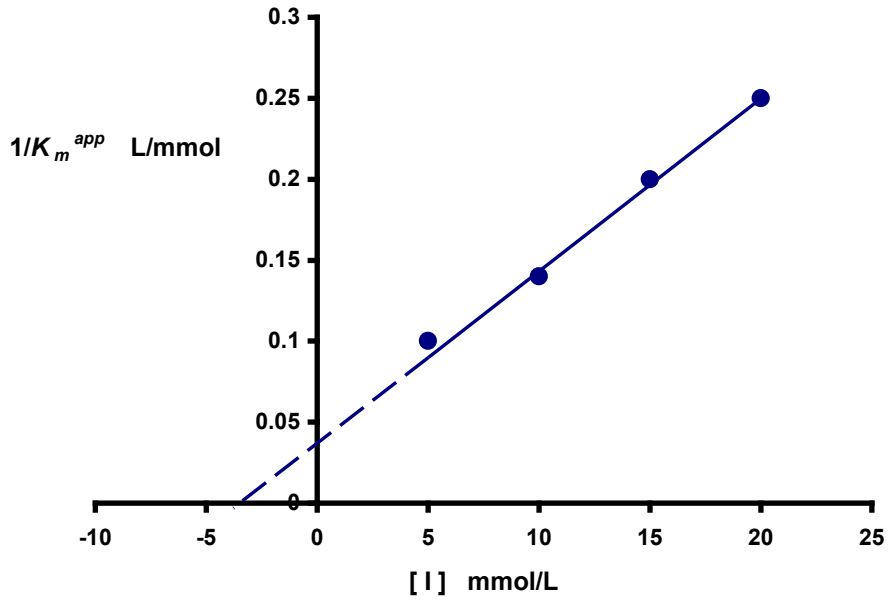
$$\frac{K_m K_i}{K_m} = -[I]$$

Cancelling K_m:

$$K_i = -[I]$$

Therefore the intercept on the [I] axis is $-K_i$

[I] (mmol/L)	5	10	15	20
$1/K_m^{app}$ (L/mmol)	0.1	0.14	0.2	0.25



When $1/K_m app = 0$, [I] = $- K_i$

From graph, when $K_m app = 0$, [I] = - 3.67 mmol/L

Therefore K_i = - (- 3.67) = 3.67 mmol/L = **3.7×10^{-3} mol/L** (2 sig figs)

The relationship between V_{max}^{app} and K_i for an uncompetitive inhibitor is:

$$V_{max}^{app} = \frac{V_{max}}{(1 + [I]/K_i)}$$

Inversion gives:

$$\frac{1}{V_{max}^{app}} = \frac{(1 + [I]/K_i)}{V_{max}}$$

Which can also be written:

$$\frac{1}{V_{max}^{app}} = \left(\frac{1}{K_i V_{max}} \times [I] \right) + \frac{1}{V_{max}}$$

When $1/V_{max}^{app} = 0$:

$$0 = \left(\frac{1}{K_i V_{max}} \times [I] \right) + \frac{1}{V_{max}}$$

Which can be rearranged to:

$$- \frac{1}{V_{max}} = \frac{[I]}{K_i V_{max}}$$

$$- \frac{K_i V_{max}}{V_{max}} = [I]$$

Cancelling V_{max} and changing the sign on both sides gives:

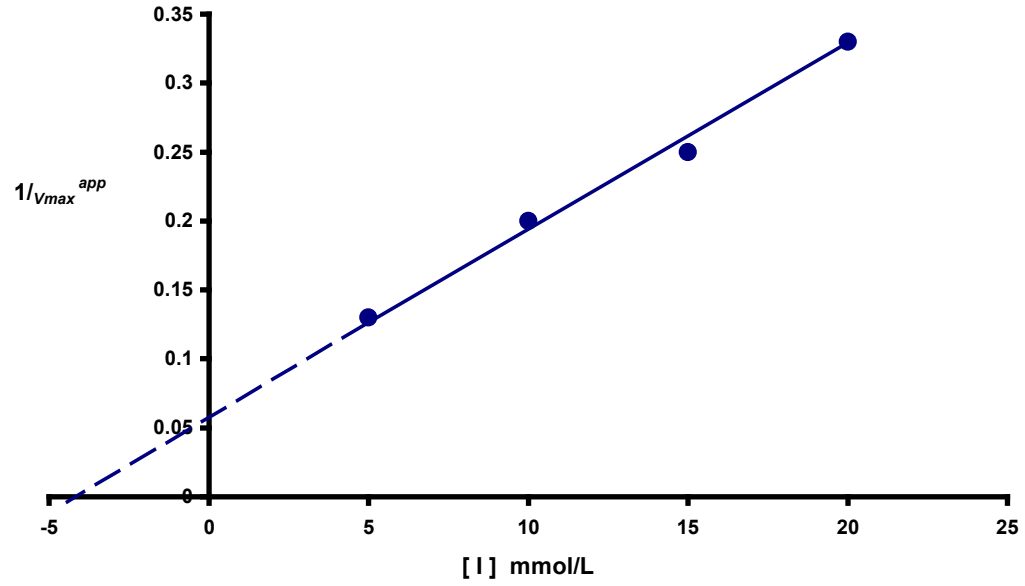
$$K_i = - [I]$$

Therefore the intercept on the [I] axis is $-K_i$.

Calculating $1/V_{max}$:

[I] mmol/L:	5	10	15	20
$1/V_{max}^{app}$	0.13	0.20	0.25	0.33

Then plotting $1/V_{max}$ versus [I]:



When $1/V_{max}^{app} = 0$, [I] = $-K_i$.

From graph, when $1/V_{max}^{app} = 0$, [I] = - 4.59 mmol/L

Therefore K_i = 4.59 mmol/L = **4.6 x 10^{-3} mol/L** (2 sig figs)

The K_i s from the two plots do not agree exactly but the value is approximately **4 x 10^{-3} mol/L**. This is due to errors inherent in manually constructing the plots and reading off the values of the intercepts.

Chapter 10

1. *The following results were obtained for a QC sample:*

   Total protein (g/L): 70, 68, 71, 65, 68, 70, 73, 69, 75, 74, 69, 71

   Calculate the mean, variance, standard deviation, coefficient of variation and 95 per cent confidence limits.

   Construct a table with columns for protein result (x) and x^2, then obtain the sum of the results in each column:

Result (x)	x^2
70	4900
68	4624
71	5041
65	4225
68	4624
70	4900
73	5329
69	4761
75	5625
74	5476
69	4761
71	5041
Total: Σx = 843	Σx^2 = 59307

Number of values of x (n) = 12

$$\text{Mean } (m) = \frac{\Sigma x}{n} = \frac{843}{12} = \textbf{70.25 g/L}$$

$$\text{Variance } (s^2) = \frac{\Sigma (x - m)^2}{n - 1}$$

As a short cut use the identity:

$$\Sigma(x - m)^2 = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$$

$$= 59{,}307 - \frac{843^2}{12}$$

$$= 59{,}307 - 59{,}221$$

$$= 86 \text{ g/L}$$

Therefore, variance $(s^2) = \frac{86}{(12 - 1)} = \frac{86}{11} =$ **7.82 g/L**

Standard deviation $(s) = \sqrt{s^2} = \sqrt{7.82} =$ **2.80 g/L**

Coefficient of variation $(CV) = \frac{s \times 100}{m} = \frac{2.80 \times 100}{70.25} =$ **4.0%**

The confidence limits of the mean are:

Mean - $(z \times s)$ to mean + $(z \times s)$

For 95% confidence limits $z = 1.96$ so that this expression becomes:

Mean - $(1.96 \times s)$ to mean + $(1.96 \times s)$

Substituting mean = 70.25 g/L and $s = 2.80$ g/L gives the 95% confidence limits:

70.25 - (1.96 x 2.80) to 70.25 + (1.96 x 2.80)

= 70.25 - 5.49 to 70.25 + 5.49

= **64.8 to 75.7 g/L** (3 sig figs)

2. *Serum thyroxine was measured in 10,000 healthy male adults. Assuming a Gaussian distribution the normal range was calculated to be 50-150 nmol/L. How many results are expected to be above 165 nmol/L?*

Assume that the normal range is the mean ± 2 standard deviations.

The mean will be the mean of the upper and lower limits:

$$m = \frac{(50 + 150)}{2} = \frac{200}{2} = 100 \text{ nmol/L}$$

The upper and lower reference limits will span 4 standard deviations:

$$s = \frac{(150 - 50)}{4} = \frac{100}{4} = 25 \text{ nmol/L}$$

Next calculate the z value for a result of 165 nmol/L:

$$z = \frac{x - m}{s} = \frac{(165 - 100)}{25} = \frac{65}{25} = 2.6$$

From tables of z, the value of P when z is equal to 2.6 is 0.002. Therefore, 0.002 of results fall outside the range: mean ± 65 nmol/L and a half of these (0.001) will be greater than 165 nmol/L.

Number of results >165 nmol/L = 0.001 x 10,000

= **10 results**

3. *Calculate the least significant difference for a change in cholesterol if the intra-individual coefficient of variation for cholesterol is 4.7% and the analytical coefficient of variation, 2.4%. A patient was changed from Atorvastatin 80 mg to Rosuvastatin 40 mg and the total cholesterol fell from 6.9 to 5.9 mmol/L. Calculate the percentage change in cholesterol and state whether this is significant.*

First calculate the total variation in terms of *CV*%:

$$CV_{Total} = \sqrt{(CV_{Analytical}^2 + CV_{Intra\text{-}individual}^2)}$$

$$= \sqrt{(2.4^2 + 4.7^2)}$$

$$= \sqrt{(5.76 + 22.09)}$$

$$= \sqrt{27.85}$$

$$= 5.28\ \%$$

Calculate the SD for the initial cholesterol value (6.9 mmol/L):

$$CV\,(\%) = \frac{s \times 100}{m}$$

$$s = \frac{CV\,(\%) \times m}{100} = \frac{5.28 \times 6.9}{100} = 0.365 \text{ mmol/L}$$

For a change to be significant the overall difference must be at least 2.8s:

Least significant change = 2.8 x 0.365 = 1.022 mmol/L

Convert this to a percentage change from the initial value:

$$\% \text{ significant change} = \frac{1.022 \times 100}{6.9} = \mathbf{14.8\%}$$

Calculate the actual percentage change in the patient's result:

$$\text{Actual \% change} = \frac{(6.9 - 5.9) \times 100}{6.9}$$

$$= \frac{1.0 \times 100}{6.9}$$

$$= \mathbf{14.5\%}$$

Since this percentage change is not greater than 14.8%, the change is **not quite statistically significant** at the 5% level of probability.

4. *Your on-call laboratory service uses 30 different methods, each of which has a 1% probability of failing QC criteria during the course of a night. Assuming that QC of any method is independent of that of the other methods, what is the probability that on any one night all methods will pass the QC criteria?*

The probability of a channel failing QC is 1% = 0.01

There are only two possible outcomes - pass or fail.

Therefore the probability of a channel passing QC is 1 - 0.01 = 0.99

This problem is analogous to flipping a coin. The joint probability of two *independent* events is the product of their individual probabilities.

Thus if a coin is tossed once the probability of 'heads' is 0.5. If the coin is tossed again then the probability of it landing 'heads' on *both* occasions is 0.5 x 0.5 = 0.25. Similarly if the probability of one channel passing QC is 0.99, then the probability of two channels passing is 0.99 x 0.99 = 0.98. The chance of three different channels passing is given by 0.99 x 0.99 x 0.99 = 0.97 i.e. $(0.99)^3$.

The general rule is:

Probability of event occurring on *n* occasions =

(probability of event occurring on a single occasion)^n

Therefore the probability of 30 channels passing QC = $(0.99)^{30}$ = 0.74

If your calculator does not have the facility to calculate x^y then the result can be easily calculated using logs:

Log_{10} (probability of 30 channels passing) = 30 x Log_{10} 0.99

= 30 x 0.00436

= - 0.131

Probability of 30 channels passing = antilog (-0.131) = **0.74**

5. *You attempt to derive a reference range for TSH for an ethnic minority population. The first 10 samples give the following results:*

Result	*n*
Between 0.5 and 1.49	*5*
Between 1.5 and 2.49	*3*
Between 2.5 and 3.49	*0*
Between 3.5 and 4.49	*1*
Between 4.5 and 5.49	*1*

On the basis of these results, what range of TSH values would encompass 95% of the ethnic minority population?

There are two problems with this set of data:

1. The individual results are not given, only the number of results falling into each class interval. The easiest way to deal with this is to assume that the results fall in the middle of the range i.e. there are 5 results within the range 0.5 to 1.49 so assume there are 5 results of the mid-point value (1.0 mU/L), similarly there are 3 samples with a value of 2 mU/L. Using this approach 10 individual results are produced which can be processed in the usual way.

2. The data are obviously skewed and do not form a Gaussian distribution. This can be overcome to some extent by taking logarithms (to the base 10) of the results then calculating the mean, SD and 95% confidence limits in the usual way. Taking antilogarithms of the confidence limits then gives the reference range.

A table can be completed in the following way:

TSH result	x = $\log_{10}$ TSH result	x^2
1.0	0	0
1.0	0	0
1.0	0	0
1.0	0	0
1.0	0	0
2.0	0.301	0.0906
2.0	0.301	0.0906
2.0	0.301	0.0906
4.0	0.602	0.3624
5.0	0.699	0.4886
n = 10	$\sum x$ = 2.204	$\sum x^2$ = 1.123

$$\text{Mean} = \frac{\sum x}{n} = \frac{2.204}{10} = 0.220$$

$$s^2 = \frac{\sum x^2 - (\sum x)^2 / n}{n - 1} = \frac{1.123 - 2.204^2 / 10}{10 - 1} = 0.0708$$

$$s = \sqrt{0.0708} = 0.266$$

Alternatively the mean and *s* can be calculated directly on most modern pocket calculators. The 95% confidence are given by mean – 1.96 *s* to mean + 1.96 *s*

= 0.220 – (1.96 x 0.266) to 0.220 + (1.96 x 0.266)

= -0.301 to 0.741 (these values are logs and so do NOT have units)

Taking antilogs (to the base 10) gives the 95% confidence limits in mU TSH/L:

0.50 to 5.51 mU/L

Although the original data may have been expressed to one or two decimal places, this information has been lost by grouping the data into class intervals. Therefore it would be more correct to quote a reference range of **less then 6 mU/L**.

6. *You are required to pipette a 9ml volume and have available a 10 ml graduated pipette which has a 2%CV associated with it's delivery volume and 5 and 2 ml volumetric pipettes each of which has a 1% CV associated with their delivery volumes. What is the error of pipetting a 9 mL volume, expressed as plus/minus mL volume?*

a) using the graduated pipette
b) using the volumetric pipettes

Assume that the error is required as 95% confidence limits i.e. ± 2 *s*.

a) Using the graduated pipette:

Calculate *s* when mean = 9 mL and *CV* = 2%:

$$\text{CV (\%)} = \frac{s \times 100}{m} =$$

$$\text{Therefore: } s = \frac{CV\,(\%) \times m}{100}$$

$$s = \frac{2 \times 9}{100} = \frac{18}{100} = 0.18 \text{ mL}$$

Therefore 95% limits $= \pm 2s = \pm 2 \times 0.18 = \pm 0.36$ mL

Error = plus/minus 0.36 mL

b) Similarly calculate the error for each of the bulb pipettes:

For 5 mL bulb with $CV = 1\%$

$$s = \frac{1 \times 5}{100} = \frac{5}{100} = 0.05 \text{ mL}$$

For 2 mL bulb with $CV = 1\ \%$:

$$s = \frac{1 \times 2}{100} = \frac{2}{100} = 0.02 \text{ mL}$$

To pipette 9 mL the 5 mL bulb is used once and the 2 mL bulb twice. Calculate the overall s:

$$s_{\text{Total}} = \sqrt{(s_{5\text{mL}}^2 + s_{2\text{mL}}^2 + s_{2\text{mL}}^2)}$$

$$= \sqrt{(0.05^2 + 0.02^2 + 0.02^2)}$$

$$= \sqrt{(0.0025 + 0.0004 + 0.0004)}$$

$$= \sqrt{0.0033}$$

$$= 0.0574 \text{ mL}$$

Therefore total error (2s) $= 2 \times 0.0574 = 0.11$ mL (2 sig figs)

Error = plus/minus 0.11 mL

7. *It has been suggested that a proposed analytical goal for an analyte is that the between batch analytical coefficient of variation should not exceed one half of the "true biological" inter-individual coefficient of variation. Calculate the percentage "expansion" of the measured reference range over the true biological reference range when this analytical goal is exactly met.*

The relationship between the overall variation, analytical variation and biological variation is:

$$CV_{Total}^2 = CV_{Analytical}^2 + CV_{Biological}^2$$

Both the analytical and biological *CV*'s share the same mean.

$$CV_{Analytical} = 0.5\ CV_{Biological}$$

Substitute this value for the analytical *CV* so as to obtain the total *CV* expressed in terms of the biological *CV*:

$$CV_{Total} = \sqrt{[(0.5\ CV_{Biological})^2 + CV_{Biological}^2]}$$

$$= \sqrt{[(0.25 \times CV_{Biological}^2) + CV_{Biological}^2]}$$

$$= \sqrt{(1.25 \times CV_{Biological}^2)}$$

$$= 1.118\ CV_{Biological}$$

The reference range encompasses a span of 4 *CV*s

Therefore biological reference range spans 4 CVs and the total reference range spans $4 \times 1.118\ CV_{Biological} = 4.47\ CV_{Biological}$

Therefore the percentage expansion is:

$$\frac{(4.47\ CV_{Biological} - 4CV_{Biological}) \times 100}{4CV_{Biological}}$$

$$\frac{CV_{Biological}\ (4.47 - 4) \times 100}{4CV_{Biological}}$$

$$\frac{(4.47 - 4) \times 100}{4} = \frac{0.47 \times 100}{4} = \mathbf{11.8\%}\ \text{(3 sig figs)}$$

Chapter 11

1. *The following analytical results were obtained on the same QC sample: 109, 91, 105, 112, 90, 115, 89, 113, 93, 94. Calculate the mean, standard deviation and standard error of the mean.*

Construct a table with columns for result (x) and x^2, then obtain the sum of the results in each column:

x	x^2
109	11,881
91	8,281
105	11,025
112	12,544
90	8,100
115	13,225
89	7,921
113	12,769
93	8,649
94	8,836

Total: $\sum x = 1011$ $\sum x^2 = 103{,}231$

$n = 10$

$$\text{Mean } (m) = \frac{\sum x}{n} = \frac{1011}{10} = \mathbf{101.1}$$

$$\text{Variance } (s^2) = \frac{\sum (x - m)^2}{n - 1}$$

$$\sum (x - m)^2 = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$= 103{,}231 - \frac{1011^2}{10}$$

$$= \quad 103{,}231 \quad - \quad 102{,}212$$

$$= \quad 1019$$

$$s^2 \quad = \quad \frac{\sum(x-m)^2}{n-1} \quad = \quad \frac{1019}{(10-1)} \quad = \quad \frac{1019}{9} \quad = \quad 113.2$$

$$\text{Standard deviation } (s) \quad = \quad \sqrt{s^2} \quad = \quad \sqrt{113.2} \quad = \quad \mathbf{10.64}$$

$$\text{Standard error of the mean } (SE_m) \quad = \quad \frac{s}{\sqrt{n}} \quad = \quad \frac{10.64}{\sqrt{10}} \quad = \quad \frac{10.64}{3.16} \quad = \quad \mathbf{3.37}$$

2. *Two laboratories measured sodium in the same plasma sample ten times. One laboratory obtained a mean of 145 mmol/L with an SD of 3 mmol/L; the other obtained a mean of 147 mmol/L with an SD of 2 mmol/L. Do the laboratories differ in their bias or imprecision?*

First lab: $m_1 = 145$ mmol/L; $s_1 = 3$ mmol/L

2nd lab: $m_2 = 147$ mmol/L; $s_2 = 2$ mmol/L

$n = 10$ for each lab

To check for bias carry out a *t*-test:

$$t \quad = \quad \frac{m_1 \; - \; m_2}{\sqrt{(s_1^2/n + \; s_2^2/n)}}$$

$$= \quad \frac{145 \; - \; 147}{\sqrt{(3^2/10 \; + \; 2^2/10)}}$$

$$= \frac{-2}{\sqrt{(0.9 + 0.4)}}$$

$$= \frac{-2}{\sqrt{1.3}} = \frac{-2}{1.14} = -1.75$$

Next calculate degrees of freedom (*DF*):

$$DF = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{[(s_1^2/n_1)^2/(n_1 - 1)] + [(s_2^2/n_2)^2/(n_2 - 1)]}$$

$$= \frac{(0.9 + 0.4)^2}{0.9^2/9 + 0.4^2/9}$$

$$= \frac{1.3^2}{0.09 + 0.018}$$

$$= \frac{1.69}{0.108}$$

$$= 15.6$$

From tables when $t = 1.75$ with 16 degrees of freedom, $P = 0.10$. Therefore there is no significant difference between the means of the two set of results i.e. **no evidence of bias.**

To compare imprecision perform an *F* ratio test:

$$F = \frac{s_1^2}{s_2^2} = \frac{3^2}{2^2} = \frac{9}{4} = 2.25$$

From tables when $F = 3.18$ (with 9 degrees of freedom for both variances), $P = 0.05$. Therefore there is no significant difference between the two variances. i.e. **no evidence of difference in imprecision.**

3. Serum thyroxine was measured in 500 healthy adults. Assuming a Gaussian distribution, the normal range was calculated to be 50-150 nmol/L. What is the probability that the mean of a set of 9 results taken at random from this population is greater than 125 nmol/L?

Assume that the normal range is the mean ± 2*s*.

The mean is the average of the upper and lower reference limit:

$$\text{Mean } (m) = \frac{(50 + 150)}{2} = \frac{200}{2} = 100 \text{ nmol/L}$$

The reference limits span 4*s* units so that that *s* is a quarter of the range:

$$\text{Standard deviation } (s) = \frac{(150 - 50)}{4} = \frac{100}{4} = 25 \text{ nmol/L}$$

Standard error of the mean (SE_m) for 9 results

$$= \frac{s}{\sqrt{n}} = \frac{25}{\sqrt{9}} = \frac{25}{3} = 8.33 \text{ nmol/L}$$

Calculate *t* for 9 results with *m* = 125 nmol/L, population mean (μ) = 100 nmol/L and SE_m = 8.33 nmol/L:

$$t = \frac{m - \mu}{SE_m} = \frac{(125 - 100)}{8.33} = \frac{25}{8.33} = 3.00 \quad (DF = n - 1 = 8)$$

From tables, for *t* = 3.00 with 8 degrees of freedom *P* = approx. 0.02. Therefore 0.02 of results fall outside of the range mean ± 25 nmol/L and a half of these results, 0.01, will be greater than 125 nmol/L.

Probability of mean of 9 results being greater than 125 nmol/L = **0.01.**

4. *It is suspected that an instrument used for near patient measurement of cholesterol is showing positive bias. The following data are the results of paired analyses of samples from ten patients measured on the standard laboratory analyser (A) and the instrument under investigation (B). Assuming that the results from the main analyser are correct, is there any evidence of bias?*

A	*B*
6.8	*7.2*
4.2	*4.5*
5.0	*4.8*
5.6	*5.9*
8.5	*8.7*
2.9	*2.8*
4.8	*4.9*
7.6	*8.1*
6.5	*6.4*
5.0	*5.2*

Since these are paired samples the results should be compared using the paired *t*-test. Construct a table with the individual differences between each pair of results (d = A – B), d^2, the difference between each d and the overall mean (m_d) for all the values of d (i.e. $d - m_d$) and their squares i.e. $(d - m_d)^2$.

A	B	d	d^2	$d - m_d$	$(d - m_d)^2$
6.8	7.2	-0.4	0.16	-0.24	0.058
4.2	4.5	-0.3	0.09	–0.14	0.020
5.0	4.8	0.2	0.04	0.36	0.130
5.6	5.9	-0.3	0.09	-0.14	0.120
8.5	8.7	-0.2	0.04	-0.04	0.002
2.9	2.8	0.1	0.01	0.26	0.070
4.8	4.9	-0.1	0.01	0.06	0.004
7.6	8.1	-0.5	0.25	-0.34	0.116
6.5	6.4	0.1	0.01	0.26	0.070
5.0	5.2	-0.2	0.04	-0.04	0.002
$n = 10$		$\sum d = -1.6$	$\sum d^2 = 0.74$	$\sum(d - m_d)^2 = 0.592$	

$$\text{Mean difference } (m_d) = \frac{\sum d}{n} = \frac{-1.6}{10} = -0.16$$

$$\text{Paired } t = \frac{m_d}{s_d/\sqrt{n}}$$

$$s_d = \sqrt{[\sum(d - m_d)^2/(n-1)]}$$

$$= \sqrt{[0.592/9]}$$

$$= \sqrt{0.0658}$$

$$= 0.256$$

Use this s_d to calculate the paired t:

$$\text{Paired } t = \frac{m_d}{s_d/\sqrt{n}} = \frac{m_d \times \sqrt{n}}{s_d}$$

$$= \frac{-0.16 \times \sqrt{10}}{0.256} = \frac{-0.16 \times 3.16}{0.256} = -1.98$$

From tables, for $t = 1.98$ (degree of freedom = $n - 1 = 9$) P is greater than 0.05.

Therefore there is no significant bias between the two methods.

5. *Four laboratories in a managed network compared the performance of their serum cholesterol assays by measuring the same sample 10 times with the following results:*

Lab			
A	*B*	*C*	*D*
7.6	*7.5*	*7.0*	*7.7*
7.3	*7.6*	*7.4*	*7.8*
7.5	*7.2*	*7.7*	*7.4*
7.7	*7.5*	*7.5*	*7.5*
7.5	*7.7*	*7.4*	*7.6*
7.6	*7.4*	*7.2*	*7.5*
7.4	*7.8*	*7.5*	*7.3*
7.8	*7.5*	*7.2*	*7.8*
7.2	*7.3*	*7.5*	*7.6*
7.5	*7.4*	*7.3*	*7.6*

Is there any significant difference in bias for serum cholesterol at this concentration between the four laboratories?

Calculate n, $\sum x$, m, $\sum x^2$, $(\sum x)^2/n$ and $\sum x^2 - (\sum x)^2/n$ for each lab:

	Lab				
	A	B	C	D	
	7.6	7.5	7.0	7.7	
	7.3	7.6	7.4	7.8	
	7.5	7.2	7.7	7.4	
	7.7	7.5	7.5	7.5	
	7.5	7.7	7.4	7.6	
	7.6	7.4	7.2	7.5	
	7.4	7.8	7.5	7.3	
	7.8	7.5	7.2	7.8	
	7.2	7.3	7.5	7.6	
	7.5	7.4	7.3	7.6	
					Totals
$\sum x$	75.1	74.9	73.7	75.8	299.5
n	10	10	10	10	40
m	7.51	7.49	7.37	7.58	
$\sum x^2$	564.29	561.29	543.53	574.8	2243.91
$(\sum x)^2/n$	564.001	561.001	543.69	574.564	2242.735
$\sum x^2 - (\sum x)^2/n$	0.289	0.289	-0.16	0.236	

Number of groups (u) = 4, number in each group (v) = 10, uv = 40

Between groups sum of squares = $\sum.(\sum x)^2/n - (\sum.\sum x)^2/uv$

= 2242.735 - $299.5^2/40$

= 2242.735 - 2242.5063

= 0.2287

Within groups sum of squares = $\sum.\sum x^2 - \sum.(\sum x)^2/n$

= 2243.91 - 2242.735

= 1.175

Total sum of squares = $\sum.\sum x^2 - (\sum.\sum x)^2/uv$

= 2243.91 - $299.5^2/40$

= 2243.91 - 2242.5063

= 1.4037

Source	**Sum of squares**	***DF***	s^2	***F***
Between groups	0.2287	3	0.0762	2.34
Within groups	1.175	36	0.0326	
Total	1.4037	39	0.0360	

From tables the probability of obtaining an F value greater than 2.84 (for 3 and 40 degrees of freedom) is 0.05. Therefore the data is homogeneous and there is **no evidence for bias between the four laboratories.**

Chapter 12

1. *Regression analysis of results using new standards (y) against old standards (x) showed a linear relationship. The regression coefficient (slope) was 1.10 and the intercept on the y axis 1.0 mmol/L. Calculate the results which would be expected using new standards for the analysis of old standards containing (a) 15 mmol/L and (b) 150 mmol/L.*

   a) Regression equation for new standards (y) upon old standards (x):

   $$y = 1.10\,x + 1.0$$

   Substitute old standard containing 15 mmol/L for *x* then solve for *y*:

   $$y = 1.10 \times 15 + 1.0$$

   $$= 16.5 + 1.0$$

   $$= \mathbf{17.5\ mmol/L}$$

   b) Substitute old standard containing 150 mmol/L for *x* then solve for *y*:

   $$y = 1.10 \times 150 + 1.0$$

   $$= 165 + 1.0$$

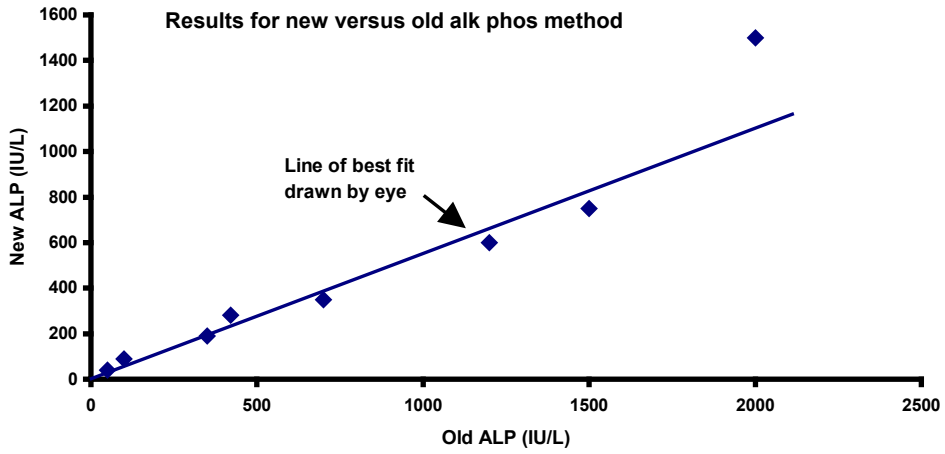
   $$= \mathbf{166\ mmol/L}$$

2. *A laboratory changed its method for the assay of serum alkaline phosphatase activity. Assay of a selection of patient's samples by both methods yielded the following data:*

ALP (Old method), IU/L:	*50*	*350*	*700*	*100*	*1500*	*2000*	*420*	*1200*
ALP (New method), IU/L:	*40*	*190*	*350*	*90*	*750*	*1500*	*280*	*600*

   A gastroenterologist has been using ALP to monitor patients on treatment. Use these data to derive an expression to convert the new ALP results to the results expected by the old method.

The first step is to check that there is a linear relationship between the two methods. This is best done by plotting the results using the new method (*y*-axis) against those obtained using the old method (*x*-axis):



The data appear to fit a straight line so linear regression analysis is appropriate.

x	y	x^2	y^2	xy
50	40	2500	1600	2000
350	190	122500	36100	66500
700	350	490000	122500	245000
100	90	10000	8100	9000
1500	750	2250000	562500	1125000
2000	1500	4000000	2250000	3000000
420	280	176400	78400	117600
1200	600	1440000	360000	720000

$\sum x = 6{,}320$ $\sum y = 3{,}800$ $\sum x^2 = 8{,}491{,}400$ $\sum y^2 = 3{,}419{,}200$ $\sum xy = 5{,}285{,}100$

$$\text{Slope of regression line } (b) = \frac{\sum(x - m_x)(y - m_y)}{\sum(x - m_x)^2}$$

$$= \frac{\sum xy \; - \; (\sum x \sum y / n)}{\sum x^2 \; - \; (\sum x)^2 / n}$$

$$b = \frac{5{,}285{,}100 - (6{,}320 \times 3{,}800/8)}{8{,}491{,}400 - (6{,}320^2/8)}$$

$$= \frac{5{,}285{,}100 - 3{,}002{,}000}{8{,}491{,}400 - 4{,}992{,}800}$$

$$= \frac{2{,}283{,}100}{3{,}498{,}600}$$

$$= 0.653 \quad (3 \text{ sig figs})$$

The value for the intercept (a) can be obtained by substituting the slope (b), the mean of x for x and the mean of y for y into the linear expression $y = bx + a$, then solving for a:

$$m_x = \frac{\sum x}{n} = \frac{6320}{8} = 790 \text{ IU/L}$$

$$m_y = \frac{\sum y}{n} = \frac{3800}{8} = 475 \text{ IU/L}$$

$$475 = (0.653 \times 790) + a$$

$$a = 475 - (0.653 \times 790)$$

$$= 475 - 516$$

$$= -41$$

Therefore regression equation of y (new results) upon x (old results):

$$\text{New method} = (\text{Old method} \times 0.65) - 41$$

Rearranging to enable easy conversion of new to old results:

$$\text{Old method} \times 0.65 = \text{New method} + 41$$

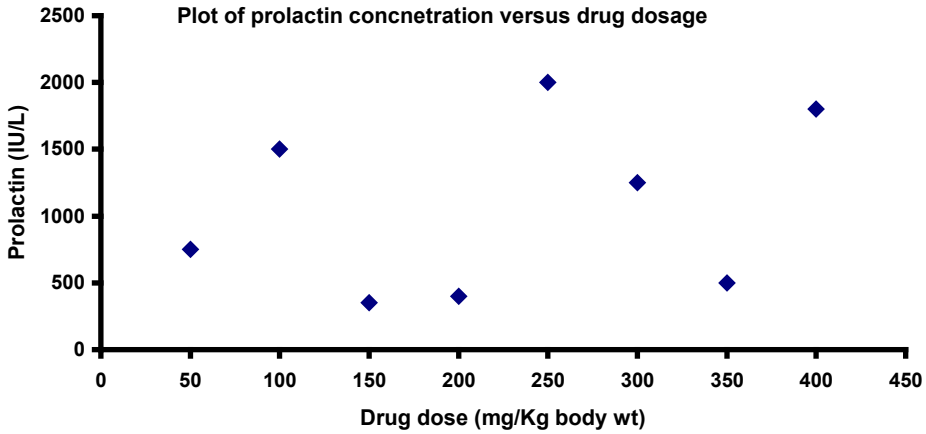
$$\textbf{Old method} = \frac{\textbf{New method} + \textbf{41}}{\textbf{0.65}}$$

3. *An endocrinologist has been using serum prolactin measurements to assess the response of patients with prolactinoma to treatment with a new drug. The following data were obtained for a series of patients:*

Drug dosage (mg/kg body wt):	*50*	*100*	*150*	*200*	*250*	*300*	*350*	*400*
Prolactin (IU/L)	*750*	*1500*	*350*	*400*	*2000*	*1250*	*500*	*1800*

Do these data show a linear relationship between drug dosage and serum prolactin concentration?

The first step is to plot the data with prolactin as the *y*-axis and drug dosage as the *x*-axis:



Visual inspection suggests that there is no significant relationship between serum prolactin concentration and drug dosage. Further evidence could be obtained by calculating the correlation coefficient:

x	x^2	y	y^2	xy
50	2500	750	562500	37500
100	10000	1500	2250000	150000
150	22500	350	122500	52500
200	40000	400	160000	80000
250	62500	2000	4000000	500000
300	90000	1250	1562500	375000
350	122500	500	250000	175000
400	160000	1800	3240000	720000

$\sum x = 1{,}800$ $\sum x^2 = 510{,}000$ $\sum y = 8{,}550$ $\sum y^2 = 12{,}147{,}500$ $\sum xy = 2{,}090{,}000$

$n = 8$

$$r = \frac{\sum xy - (\sum x \sum y/n)}{\sqrt{\{[\sum x^2 - (\sum x)^2/n]\,[\sum y^2 - (\sum y)^2/n]\}}}$$

$$= \frac{2{,}090{,}000 - (1{,}800 \times 8{,}550/8)}{\sqrt{\{[510{,}000 - 1{,}800^2/8]\,[12{,}147{,}500 - 8{,}550^2/8]\}}}$$

$$= \frac{2{,}090{,}000 - 1{,}923{,}750}{\sqrt{\{[510{,}000 - 405{,}000]\,[12{,}147{,}500 - 9{,}137{,}813]\}}}$$

$$= \frac{166{,}250}{\sqrt{\{105{,}000 \times 3{,}009{,}687\}}}$$

$$= \frac{166{,}250}{562{,}154}$$

$$= 0.30 \text{ (2 sig figs)}$$

From tables, for $r = 0.30$ with 7 degrees of freedom, $P > 0.1$. Therefore there is **no significant correlation between drug dosage and serum prolactin**.

4. *A research paper contains the following statement:*

"A good correlation was obtained when 45 patient samples were analysed by methods A and B (r = 0.90, B = 1.05A – 10)…." Comment on this statement.

No evidence is presented that the relationship between the two variables is linear.

Correlation analysis is not the best approach to comparing two analytical methods – as they both measure the same analyte it would be surprising if there were no correlation. Analysis of difference plots would be more appropriate.

The standard error of the slope (1.05) is not given.

The standard deviation of the residual (s_{res} or sy_x) is not given – this is the best indicator of the goodness of fit of the data to the regression line.

Chapter 13

1. A test for a particular disease has a sensitivity of 95% and a specificity of 95%. Calculate the predictive value of both a positive and a negative test result in a population in which the prevalence of the disease is:

a) 1 in 2
b) 1 in 5000

a) It is easiest to work with proportions rather than percentages or absolute numbers of results. The contingency table to use is:

	Positive result	**Negative result**	**Total**
Patients *with* disease	TP	FN	Prevalence
Patients *without* disease	FP	TN	1 – prevalence

If the prevalence of disease is 1 in 2 i.e. 0.5, then 1 – prevalence is also 0.5 so this table becomes:

	Positive result	**Negative result**	**Total**
Patients *with* disease	TP	FN	0.5
Patients *without* disease	FP	TN	0.5

The next task is determine values for TP, FN, FP and TN using the stated sensitivity and specificity:

$$\text{Sensitivity} = \frac{TP}{TP + FN} = 0.95$$

Substitute (TP + FN) = 0.5, then solve for TP:

$$\frac{TP}{0.5} = 0.95 \text{ so that } TP = 0.5 \times 0.95 = 0.475$$

and FN = 0.5 - TP = 0.5 - 0.475 = 0.025

Similarly using specificity:

$$\text{Specificity} = \frac{TN}{TN + FP} = 0.95$$

$$\frac{TN}{0.5} = 0.95 \quad \text{so that} \quad TN = 0.5 \times 0.95 = 0.475$$

and FP = 0.5 - TN = 0.5 - 0.475 = 0.025

Inserting these values into the contingency table gives:

	Positive result	**Negative result**	**Total**
Patients *with* disease	0.475	0.025	0.5
Patients *without* disease	0.025	0.475	0.5

These values are then used to calculate positive and negative predictive values:

$$PV(+) = \frac{TP}{TP + FP} = \frac{0.475}{0.475 + 0.025} = \mathbf{0.95\ (95\%)}$$

$$PV(-) = \frac{TN}{TN + FN} = \frac{0.475}{0.475 + 0.025} = \mathbf{0.95\ (95\%)}$$

b) With a prevalence of 1 in 5,000 (=0.0002) the contingency table becomes:

	Positive result	**Negative result**	**Total**
Patients *with* disease	TP	FN	0.0002
Patients *without* disease	FP	TN	0.9998

TP = 0.0002 x 0.95 = 0.00019

FN = 0.0002 - 0.00019 = 0.00001

TN = 0.9998 x 0.95 = 0.94981

FP = 0.9998 - 0.94981 = 0.04999

So the contingency table becomes:

	Positive result	**Negative result**	**Total**
Patients *with* disease	0.00019	0.00001	0.0002
Patients *without* disease	0.04999	0.94999	0.9998

Use these values to calculate positive and negative predictive values:

$$PV(+) = \frac{TP}{TP + FP} = \frac{0.00019}{0.00019 + 0.04999} = \mathbf{0.004\ (0.4\%)}$$

$$PV(-) = \frac{TN}{TN + FN} = \frac{0.94999}{0.94999 + 0.00001} = \mathbf{1.00\ (100\%)}$$

2. *The table shows data from two urinary screening tests for the detection of phaeochromocytoma.*

Test	*Sensitivity*	*Specificity*
VMA	*96.7%*	*99.1%*
Total metanephrines	*100%*	*98%*

Both tests were used to screen a population of 100,000 hypertensive patients in which the incidence of phaeochromocytoma is known to be 0.5%.

a) How many patients with phaeochromocytoma were missed by the VMA test?

b) How many patients were incorrectly diagnosed as having phaeochromocytoma using the metanephrine test?

c) Which test would you use to screen a hypertensive population for phaeochromocytoma? Give reasons for your choice.

a) The number of patients with phaeochromocytoma missed by the VMA test is the number of *false negatives* using this test.

First calculate the proportion of false negatives i.e. use the sensitivity expressed as a proportion (0.967) rather than percentage (96.7%) and the prevalence calculated as follows:

$$\text{Prevalence} = \frac{0.5}{100} = 0.005$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

Substitute (TP + FN) = prevalence = 0.005, and sensitivity = 0.967 and solve for TP:

$$\text{Sensitivity} = \frac{TP}{0.005} = 0.967$$

$$TP = 0.967 \times 0.005 = 0.004835$$

Since $TP + FN = 0.005$

$$FN = 0.005 - 0.004835 = 0.000165$$

Multiply this *proportion* by the total number screened to obtain the *number* of false negatives (i.e. cases of phaeochromocytoma missed):

$$\text{Patients missed} = 0.000165 \times 100{,}000$$

$$= \mathbf{16.5} \text{ (2 sig figs)}$$

b) The proportion of patients incorrectly diagnosed with phaeochromocytoma using the metanephrine test is the proportion of false positives which can be calculated from the specificity and prevalence:

$$\text{Specificity} = \frac{TN}{TN + FP} = 0.98$$

$$TN + FP = 1 - \text{prevalence} = 1 - 0.005 = 0.995$$

$$\text{Specificity} = \frac{TN}{0.995} = 0.98$$

$$TN = 0.98 \times 0.995 = 0.9751$$

$$FP = (1 - \text{prevalence}) - TN$$

$$FP = 0.995 - 0.9751 = 0.0199$$

Multiply the proportion of false positives by the total number tested to give the absolute number of false positives i.e. the number of patients incorrectly diagnosed with phaeochromocytoma by the metanephrine test:

Number incorrectly diagnosed = 0.0199 x 100,000

= **1990**

c) Probably the best way to decide which is the best test is to calculate the positive and negative predictive values for each test:

For VMA:

	Positive result	**Negative result**	**Total**
Patients *with* disease	0.004835	0.00165	0.005
Patients *without* disease	0.008955	0.986045	0.995

$$PV(+) = \frac{TP}{TP + FP} = \frac{0.004835}{0.004835 + 0.008955} = \frac{0.004835}{0.01379} = \mathbf{0.35}$$

$$PV(-) = \frac{TN}{TN + FN} = \frac{0.986045}{0.986045 + 0.00165} = \frac{0.986045}{0.987695} = \mathbf{0.998}$$

For metanephrines:

	Positive result	**Negative result**	**Total**
Patients *with* disease	0.005	0.000	0.005
Patients *without* disease	0.0199	0.9751	0.995

$$PV(+) = \frac{TP}{TP + FP} = \frac{0.005}{0.005 + 0.0199} = \frac{0.005}{0.0249} = \mathbf{0.20}$$

$$PV(-) = \frac{TN}{TN + FN} = \frac{0.9751}{0.9751 + 0.000} = \frac{0.9751}{0.9751} = 1.00$$

To summarize:

Test	PV(+)	PV(-)
VMA	0.35	0.998
Metanephrines	0.20	1.00

Although the VMA test produces less false positives (i.e. higher PV+) this is achieved at the expense of missing approximately 1 in 3 (FN/prevalence = 0.33) patients with phaeochromocytoma. Although the phaeochromocytoma produces more false positives (i.e. lower PV+) this is achieved without missing any cases of phaeochromocytoma (i.e. no false negatives). **On balance total metanephrines is the better test.**

3. *A new laboratory test has a sensitivity of 85% and a specificity of 90%. The incidence of disease in a population considered at risk is 0.10. What is the predictive value of*

a) a positive result?
b) a negative result?

Start by drawing up a contingency table:

	Positive result	**Negative result**	**Total**
Patients *with* disease	TP (Sens x prev)	FN (Prev – TP)	Prev (TP + FN)
Patients *without* disease	FP [(1 – prev) – TN]	TN {Spec x (1 – prev)]	1 – prev (FP + TN)
Total	TP + FP	TN + FN	1

Using sensitivity and specificity expressed as proportions instead of percentages i.e. sensitivity = 0.85 and specificity = 0.90 fill in the above table:

	Positive result	**Negative result**	**Total**
Patients *with* disease	0.085 (Sens x prev)	0.015 (Prev – TP)	0.10 (TP + FN)
Patients *without* disease	0.09 [(1 – prev) – TN]	0.81 {Spec x (1 – prev)]	0.90 (FP + TN)
Total	0.175 TP + FP	0.825 TN + FN	1

a) Predictive value of a positive result:

$$PV(+) = \frac{TP}{TP + FP} = \frac{0.085}{0.175} = \mathbf{0.49}\ (2\ \text{sig figs})$$

b) Predictive value of a negative result:

$$PV(-) = \frac{TN}{TN + FN} = \frac{0.81}{0.825} = \mathbf{0.98}\ (2\ \text{sig figs})$$

4. *A proposed diagnostic serological test for coeliac disease was evaluated in 200 consecutive patients referred to a paediatric gastroenterology service in whom the condition was suspected clinically. The test result was compared with the diagnosis as established by biopsy, withdrawal of gluten and response to re-challenge. On this basis 76 children had the condition of whom only 64 gave a positive test result: 10 positive test results occurred in children who were shown not to have coeliac disease. Calculate the sensitivity and specificity of the test and the predictive value of a positive result.*

The prevalence of disease amongst the population of 200 is 76, so that (1 – prevalence) = 200 – 76 = 124

64 of these 76 gave a positive test result = true positives

10 were positive amongst those without celiac disease = false positives

Set up a 2 x 2 contingency table for these results:

	Positive result	**Negative result**	**Total**
Patients *with* disease	TP (Sens x prev)	FN (Prev – TP)	Prev (TP + FN)
Patients *without* disease	FP [(1 – prev) – TN]	TN {Spec x (1 – prev)]	1 – prev (FP + TN)
Total	TP + FP	TN + FN	1

Fill in this table working with the data given in the question:

	Positive result	**Negative result**	**Total**
Patients *with* disease	64	12	76
Patients *without* disease	10	114	124
Total	74	126	200

N.B. In a question like this when the sensitivity and specificity is not given and actual numbers of results are supplied it is probably easiest to work with absolute numbers rather than proportions.

$$\text{Sensitivity} = \frac{TP}{TP + FN} = \frac{64}{76} = \mathbf{0.84} \text{ (or 84%) 2 sig figs}$$

$$\text{Specificity} = \frac{TN}{TN + FP} = \frac{114}{124} = \mathbf{0.92} \text{ (or 92%) 2 sig fgs}$$

$$\text{PV(+)} = \frac{TP}{TP + FP} = \frac{64}{74} = \mathbf{0.86} \text{ (or 86%) 2 sig figs}$$

5. *In a cancer clinic where the prevalence of ovarian malignancy is 40%, a tumour marker has a specificity of 88% and a sensitivity of 92%. Calculate the predictive value of a positive test result. If this test was used as a screening tool in all patients attending a general gynaecological clinic with a cancer prevalence of 0.4%, what would be the predictive value of a positive test in this population?*

Set up a 2 x 2 contingency table then fill in the gaps using a prevalence of 0.4, sensitivity of 0.92 and specificity of 0.88:

	Positive result	**Negative result**	**Total**
Patients *with* disease	TP (Sens x prev)	FN (Prev – TP)	Prev (TP + FN)
Patients *without* disease	FP [(1 – prev) – TN]	TN {Spec x (1 – prev)]	1 – prev (FP + TN)
Total	TP + FP	TN + FN	1

	Positive result	**Negative result**	**Total**
Patients *with* disease	0.368 (Sens x prev)	0.032 (Prev – TP)	0.4 (TP + FN)
Patients *without* disease	0.072 [(1 – prev) – TN]	0.528 {Spec x (1 – prev)]	0.6 (FP + TN)
Total	0.44	0.56	1

$$PV(+) = \frac{TP}{TP + FP} = \frac{0.368}{0.44} =$$ **0.84** or 84% (2 sig figs)

Recalculate the above table using a prevalence of 0.4 % (i.e. 0.004):

	Positive result	**Negative result**	**Total**
Patients *with* disease	0.00368 (Sens x prev)	0.00032 (Prev – TP)	0.004 (TP + FN)
Patients *without* disease	0.12 [(1 – prev) – TN]	0.876 {Spec x (1 – prev)]	0.996 (FP + TN)
Total	0.12368	0.87632	1

$$PV(+) = \frac{TP}{TP + FP} = \frac{0.00368}{0.12368} =$$ **0.030** or 3.0 % (2 sig figs)

6. *A certain disease has a prevalence of 5 percent. A diagnostic test was applied to a random sample of 400 individuals from this population and yielded 15 true positive and 30 false positive results. Calculate: a) the positive predictive value of the test applied to this population, b) the pre-test odds of disease, c) the likelihood ratio positive; d) the post test odds of disease for a positive result, and e) the post-test probability of disease for a positive result.*

The 2 x 2 contingency table can be set up as follows:

	Positive result	**Negative result**	**Total**
Patients *with* disease	TP (Sens x prev)	FN (Prev – TP)	Prev (TP + FN)
Patients *without* disease	FP [(1 – prev) – TN]	TN {Spec x (1 – prev)]	1 – prev (FP + TN)
Total	TP + FP	TN + FN	1

Complete the table using a prevalence of 5% = 0.05 which with a total of 400 individuals gives a prevalence in absolute numbers of 0.05 x 400 = 20.

	Positive result	**Negative result**	**Total**
Patients *with* disease	15 (Sens x prev)	5 (Prev – TP)	20 (TP + FN)
Patients *without* disease	30 [(1 – prev) – TN]	350 {Spec x (1 – prev)]	380 (FP + TN)
Total	45	355	400

a) $PV(+) = \frac{TP}{TP + FP} = \frac{15}{45} =$ **0.33** or 33% (2 sig figs)

b) Pre-test odds $= \frac{\text{Prevalence}}{1 - \text{prevalence}}$

$= \frac{20}{380}$

$= 0.053$ (2 sig figs)

c) $LR+ = \frac{\text{probability of +ve test with disease}}{\text{probability of a +ve test without disease}} = \frac{\text{sensitivity}}{(1 - \text{specificity})}$

Sensitivity $= \frac{TP}{TP + FN} = \frac{15}{20} = 0.75$

Specificity $= \frac{TN}{TN + FP} = \frac{350}{380} = 0.92$ (2 sig figs)

$LR+ = \frac{0.75}{(1 - 0.92)} = \frac{0.75}{0.08} =$ **9.4** (2 sig figs)

d) Post-test odds = Pre-test odds x LR+

$= 0.053 \times 9.4$

$=$ **0.50** (2 sig figs)

e) Post-test probability $= \frac{\text{Post-test odds}}{(1 + \text{post-test odds})}$

$= \frac{0.50}{(1 + 0.50)}$

$= \frac{0.50}{1.50}$

$=$ **0.33**

7. *A two-stage sequential test strategy is used to screen for a rare inherited disease. The prevalence of the disease is 0.0005. The initial test has a sensitivity of 98% and specificity of 95%, the follow-up test a sensitivity of 95% and specificity of 99%. What is the probability of a patient with a positive result for the follow-up test having the disease?*

If the prevalence of disease is 0.0005 then the pre-test odds can be calculated:

$$\text{Pre-test odds} = \frac{\text{Prevalence}}{(1 - \text{prevalence})} = \frac{0.0005}{(1 - 0.0005)} = \frac{0.0005}{0.9995} = 0.000500$$

$$\text{LR+} = \frac{\text{probability of +ve test with disease}}{\text{probability of a +ve test without disease}} = \frac{\text{sensitivity}}{(1 - \text{specificity})}$$

$$\text{For the 1st test: LR+} = \frac{0.98}{(1 - 0.95)} = \frac{0.98}{0.05} = 19.6$$

$$\text{For the 2nd test: LR+} = \frac{0.95}{(1 - 0.99)} = \frac{0.95}{0.01} = 95$$

Post test odds =

Pre-test odds x likelihood ratio (1st test) x likelihood ratio (2nd test)

$$\text{Post test odds} = 0.000500 \times 19.6 \times 95 = 0.931$$

$$\text{Post-test probability} = \frac{\text{Post-test odds}}{(1 + \text{post test odds})}$$

$$= \frac{0.931}{1.931}$$

$$= \mathbf{0.48} \text{ or } 48\% \text{ (2 sig figs)}$$

Chapter 14

1. *A study into the effect of nutritional supplements on patients with Crohn's disease involved measuring serum albumin both before and after supplementation for a four week period. During this period the mean serum albumin level increased from 25 g/L to 30 g/L. The study involved 40 patients with a standard deviation for albumin concentration of 10 g/L. What is the power of this study to detect a 5 g/L change in serum albumin at the 5% level of probability?*

   The power can be calculated from the following expression:

$$z_\alpha + z_\beta = \frac{\Delta \sqrt{n}}{s}$$

   Δ = difference between the means of the two groups = 30-25 = 5 g/L

   n = number of subjects in the study = 40

   s = standard deviation = 10 g/L

   Substitute these values to obtain $z_\alpha + z_\beta$

$$z_\alpha + z_\beta = \frac{5 \sqrt{40}}{10} = \frac{5 \text{ x } 6.32}{10} = 3.16$$

   Since the probability (P) used as a decision level is 0.05, the corresponding z value (obtainable from tables) is 1.96 (the question only requires detection of a change – which could be either positive or negative – so both sides of the distribution are being used). Therefore, α = 0.05 and z_α is 1.96. Substitute this value for z_α and solve for z_β:

$$z_\beta = 3.16 - z_\alpha = 3.16 - 1.96 = 1.20$$

   From tables of z, the value for β (i.e. proportion of area under the curve) corresponding a to z_β of 1.20 is 0.1151 (single sided probability).

   Therefore power = $(1 - \beta)$ = (1 - 0.1151) = **0.88** or 88% (2 sig figs)

2. *It is proposed to set up a study to determine the effect of dietary modification on serum cholesterol. The population to be studied has a mean serum cholesterol of 7.5 mmol/L with standard deviation of 2.5 mmol/L. What number of participants need to be recruited in order to demonstrate a lowering of serum cholesterol by 10% (using alpha = 0.05 as a critical value) with a power of 90%?*

The expression for calculating sample size is:

$$n = [s\,(z_\alpha + z_\beta) / \Delta]^2$$

s = standard deviation = 2.5 mmol/L

Δ = difference between the means

= Final cholesterol - Initial cholesterol

= (90% x 7.5) - 7.5 (since cholesterol is required to be lowered by 10%)

= 6.75 - 7.5

= - 0.75 mmol/L

The required power is 90%

Therefore $(1 - \beta) = 0.9$ and $\beta = 1 - 0.9 = 0.1$

From tables the corresponding z value (i.e. z_β) is 1.28 (one sided value).

The decision level used is a probability of 0.05 (α) with a corresponding z value for one side of the distribution (since we are required to detect a decrease in cholesterol) (z_α) of 1.64.

Substitute these values and solve for n:

$$n = [2.5\,(1.64 + 1.28) / -0.75]^2$$

$$= [2.5 \times 2.92 / -0.75]^2$$

$$= 9.73^2$$

= **95** (2 sig figs)

Chapter 15

1. *A 0.5 mL sample of urine is extracted into dichloromethane. An aliquot of the extract is analysed by HPLC and found to give an apparent original concentration of 320 nmol/L of analyte Y. 100 μL of Y standard with a concentration of 880 nmol/L is added to a further 0.5 mL sample of the same urine and the sample mixed. 0.5 mL of the mixed sample is then processed as before, giving a measured concentration of 405 nmol/L. Calculate the recovery of analyte Y.*

$$\text{Recovery \%} = \frac{\text{Increase in concentration upon adding standard} \times 100}{\text{Concentration of standard added}}$$

Allowance must be made for dilution of both the sample and standard when they are mixed – since only 0.5 mL of the mixture is used for the assay.

Concentration of Y from urine in the mixture

$$= \frac{\text{Initial concentration} \times \text{Volume of urine (mL)}}{\text{Volume of mixture (mL)}}$$

Since initial concentration = 320 nmol/L
Mixture = 0.5 mL urine + 0.1 mL standard = 0.6 mL

$$\text{Concentration of Y from urine} = \frac{320 \times 0.5}{0.6} = 266.7 \text{ nmol/L}$$

$$\text{Similarly concentration of standard in mixture} = \frac{880 \times 0.1}{0.6} = 146.7 \text{ nmol/L}$$

$$\text{Recovery (\%)} = \frac{(\text{Measured concn} - \text{concn from urine}) \times 100}{\text{Standard added}}$$

$$= \frac{(405 - 266.7) \times 100}{146.7}$$

$$= \frac{138.3 \times 100}{146.7}$$

$$= \mathbf{94\%} \quad (2 \text{ sig figs})$$

2. *A new method for HCG in urine is being evaluated. The concentration in a sample from a pregnant woman is measured at 8240 IU/L. A 50 µL aliquot of an international standard containing 50,000 IU/L is added to 450 µL of the same urine sample and the sample mixed. On measuring the mixed sample, the new concentration is found to be 12100 IU/L. What is the recovery of HCG by this method?*

Calculate the expected concentrations in the mixture from the urine and the standard separately:

$$\text{Urine HCG in mixture} = \frac{8240 \times 450}{500} = 7{,}416 \text{ IU/L}$$

$$\text{Standard HCG in mixture} = \frac{50{,}000 \times 50}{500} = 5{,}000 \text{ IU/L}$$

$$\%\text{ recovery} = \frac{\text{HCG recovered} \times 100}{\text{HCG added}}$$

$$= \frac{(\text{Measured HCG in mixture} - \text{Expected HCG from urine}) \times 100}{\text{HCG added}}$$

$$= \frac{(12{,}100 - 7{,}416) \times 100}{5{,}000}$$

$$= \frac{4{,}684 \times 100}{5{,}000}$$

$$= \mathbf{94\%} \quad (2 \text{ sig figs})$$

3. *Measurement of plasma AFP is used to monitor a patient with a teratoma. If the initial concentration was 10,200 U/L what plasma level would you expect to find 21 days after successful surgery? Assume the half-life of AFP is 5.5 days.*

Assuming the clearance of AFP follows first-order kinetics the rate equation is:

$$\ln Cp_t = \ln Cp_0 - k_d.t$$

Cp_t = concentration of AFP after 21 days = Cp_{21}
Cp_0 = initial concentration of AFP = 10,200 U/L
t = time period = 21 days

k_d = elimination rate constant which can be calculated from the half-life ($t_{½}$):

$$k_d = \frac{0.693}{t_{½}} = \frac{0.693}{5.5} = 0.126 \text{ days}^{-1}$$

Substitute these values and solve for Cp_{21}:

$$\ln Cp_{21} = \ln 10{,}200 - 0.126 \times 21$$

$$\ln Cp_{21} = 9.230 - 2.646 = 6.584$$

$$Cp_{21} = \text{antilog}_e\ 6.584 = \mathbf{723\ U/L}$$

4. *A radioisotope has a half-life of 21 days. How long will it take for the activity to fall to 10% of the initial value?*

The decay of a radioisotope follows first-order kinetics:

$$\ln A_t = \ln A_0 - k_d.t$$

A_t = activity at time t = 0.1
A_0 = initial activity = 1
t = time for activity to fall to 10% of initial value = ?

k_d = decay constant which can be calculated from the given half-life ($t_{½}$):

$$k_d = \frac{0.693}{t_{½}} = \frac{0.693}{21} = 0.033 \text{ days}^{-1}$$

Substitute these values and solve for t:

$$\ln 0.1 = \ln 1 - 0.033.t$$

$$-2.303 = 0 - 0.033.t$$

$$0.033.t = 2.303$$

$$t = \frac{2.303}{0.033}$$

$$= \textbf{70 days} \text{ (2 sig figs)}$$

An alternative approach is to use the expression:

$$\log_{10} AR = -0.30.N$$

AR = ratio of final to initial activity = 0.1
N = number of half-lives for this change to occur

Therefore $\log_{10} 0.1 = -0.30.N$

$$N = \frac{1}{0.30} = 3.333 \text{ half-lives}$$

As $t_{½}$ = 21 days, $t = 3.333 \times 21 = 70$ days

5. *In normal pregnancy serum beta hCG has a doubling time of approximately 2 days. How long will it take for the serum level to increase ten-fold?*

Exponential growth obeys the first-order rate equation:

$$\ln Cp_t = \ln Cp_0 + k_d.t$$

If we take 1 as the initial concentration then a 10-fold increase will result in a concentration of 10.

Cp_t = concentration after time t = 10
Cp_0 = initial concentration = 1
k_d = specific growth rate, which can be calculated from the doubling time (t_d):

$$k_d = \frac{0.693}{t_d} = \frac{0.693}{2} = 0.3465 \text{ day}^{-1}$$

t = time taken for concentration to increase 10-fold = ?

Substitute these values and solve for t:

$$\ln 10 = \ln 1 + 0.3465.t$$

$$2.303 = 0 + 0.3465.t$$

$$0.3465.t = 2.303$$

$$t = \frac{2.303}{0.3465} = \textbf{6.6 days} \text{ (2 sig figs)}$$

Alternatively the following expression can be used:

$$\log_{10} CR = 0.30\ N$$

where CR is the concentration ratio = 10:1

N = number of doubling times required to achieve this ratio

Substitute $CR = 10$ and solve for N:

$$\log_{10} 10 = 0.30\ N$$

$$1 = 0.30\ N$$

$$N = \frac{1}{0.30} = 3.333$$

Therefore time taken = N x t_d

= 3.333 x 2

= 6.7 days

6. *A patient receiving parenteral nutrition is receiving 11.8 g nitrogen/24 h as amino acids. Urinary urea excretion is 580 mmol/24 h. Indicating what assumptions you make, calculate whether she is in positive or negative nitrogen balance.*

$$\text{Nitrogen excretion(g/24h)} = \frac{\text{Urea excretion (mmol/24 h) x 28}}{1{,}000}$$

$$= \frac{580 \text{ x } 28}{1{,}000}$$

= 16.24 g/24 g

Nitrogen balance (g/24 h) = Nitrogen intake (g/24 h) - Nitrogen excretion (g/24 h)

= 11.8 - 16.24

= - 4.44 g/24h

If 20% is added to the urinary excretion to allow for other urinary losses and a further 2 g/day added to allow for losses by other routes then the nitrogen excretion becomes:

Corrected nitrogen excretion = [Urea nitrogen excretion (g/24 h) x 1.2] + 2.0

= (16.24 x 1.2) + 2.0

= 19.49 + 2.0

= 21.49 g/24 h

and the corrected nitrogen balance becomes:

Corrected nitrogen balance (g/24 h) = 11.8 - 21.49

= **- 9.7 g (Negative balance)**

7. *A 30 min basal gastric secretion sample (total volume 27 mL) required 2.5 mL of 0.1 M NaOH to titrate 5 mL of the material to pH 7.4. Calculate the basal acid secretion rate in mmol/h.*

$$M_1 \times V_1 = M_2 \times V_2$$

M_1 = molar concentration of HCl in gastric fluid = ?
V_1 = volume of gastric fluid used in titration = 5 mL
M_2 = molar concentration of NaOH = 0.1 M
V_2 = titre of NaOH = 2.5 mL

$$M_1 \times 5 = 0.1 \times 2.5$$

$$M_1 = \frac{0.1 \times 2.5}{5}$$

$$= 0.05 \text{ M}$$

Since the answer is required in mmol multiply by 1,000:

$$\text{HCl concentration} = 0.05 \times 1{,}000 = 50 \text{ mmol HCl/L gastric fluid}$$

Divide by 1,000 to give the acid output per mL of gastric fluid then mutiply by the total volume of gastric fluid collected (27 mL) to obtain the total output of acid:

$$\text{Total HCl output} = \frac{50 \times 27}{1{,}000} = 1.35 \text{ mol HCl /27 mL gastric fluid}$$

Since the gastric fluid was collected over 30 min, multiply this result by 2 to obtain the amount of HCl secreted in 1 h:

$$\text{Rate of HCl excretion} = 1.35 \times 2$$

$$= \mathbf{2.7 \text{ mmol/h}}$$

8. *A five-day faecal fat collection was homogenised and diluted to 1500 mL. A 10 mL aliquot of the homogenate was subjected to hydrolysis and the fatty acids were extracted. The volume of 0.05 M sodium hydroxide required to effect neutralisation was 48 mL. Calculate the fat excretion in mmol/24 h.*

First calculate the fatty acid concentration in the homogenate:

$$M_1 \times V_1 = M_2 \times V_2$$

M_1 = molar concentration of fatty acids in homogenate = ?
V_1 = volume of homogenate titrated = 10 mL
M_2 = molar concentration of NaOH used in titration = 0.05 M
V_2 = titre of 0.05 M NaOH = 48 mL

$$M_1 \times 10 = 0.05 \times 48$$

$$M_1 = \frac{0.05 \times 48}{10} = 0.24 \text{ mol/L}$$

Multiply by 1,000 to convert this concentration to mmol/L:

$$\text{Fatty acid concentration} = 0.24 \times 1{,}000$$

$$= 240 \text{ mmol fatty acid/L homogenate}$$

Multiply by the total volume (in litres) of the homogenate to obtain the total fatty acid output over the 5-day collection period:

$$\text{Fatty acid output} = 240 \times 1.50 = 360 \text{ mmol fatty acid/5 days}$$

Division by 5 gives the daily fatty acid output:

$$\text{Daily fatty acid output} = \frac{360}{5} = 72 \text{ mmol fatty acid/24 h}$$

Assuming that all the fatty acids were liberated from triglyceride then division by 3 gives the total fat output:

$$\text{Fat output} = \frac{72}{3} = \textbf{24 mmol fat/24 h (as triglyceride)}$$

9. *Gas chromatography for a drug involves adding equal amounts of internal standard to standard or sample prior to analysis. The following peak areas were obtained:*

Sample	*Peak area*	
	Internal standard	*Drug*
Standard (200 nmol/L)	*50,000*	*200,000*
Patient	*40,000*	*150,000*

Calculate the drug concentration in the sample.

Divide the drug peak area by the internal standard peak area to give the peak height ratio (PHR) for both standard and patient:

Sample	Peak area		
	Internal standard	Drug	PHR
Standard (200 nmol/L)	50,000	200,000	4.00
Patient	40,000	150,000	3.75

Assuming that the PHR is directly proportional to concentration then concentration of the drug in the patient sample can be calculated from the relationship:

$$\frac{PHR_{Patient}}{Concentration_{Patient}} = \frac{PHR_{Standard}}{Concentration_{Standard}}$$

Substitute the PHR values and standard concentration to obtain the drug concentration in the patient sample:

$$\frac{3.75}{Concentration_{Patient}} = \frac{4.00}{200}$$

$$Concentration_{Patient} = \frac{3.75 \times 200}{4.00}$$

$$= \mathbf{188\ nmol/L} \quad (3 \text{ sig figs})$$

10. Genotyping of a group of 100 unrelated individuals for a two-allele polymorphism showed that the allele frequencies were:

A 0.65
B 0.35

Calculate the expected percentages of heterozygotes (AB) and homozygotes (AA and BB) in the group.

Frequency of allele A = p = 0.65

Frequency of allele B = q = 0.35

The possible combinations are AA, AB and BB.

If the conditions for the Hardy-Weinberg equilibrium are met then the frequencies of the three genotypes are:

AA = p^2 = 0.65^2 = 0.4225

AB = $2pq$ = 2 x 0.65 x 0.35 = 0.455

BB = q^2 = 0.35^2 = 0.1225

Therefore % heterozygotes (AB) = 0.455 x 100 = **45.5%**

and % homozygotes (AA and BB) = (0.4225 + 0.1225) x 100 = **54.5%**

11. The prevalence of an inherited metabolic disease (inherited in an autosomal recessive manner due to a single allele) is 1 in 2,500. A survey identified 1 in 50 of the population as asymptomatic carriers. Is this finding consistent with a population in a Hardy-Weinberg equilibrium?

Let the dominant gene be *A* and the recessive gene *a*. As the inheritance of the disease is autosomal recessive only the homozygous recessive genotype (*aa*) expresses the disease.

The incidence of the recessive disorder (*aa*) = 1 in 2,500 = $\frac{1}{2{,}500}$ = 0.00040

Incidence of carriers (*Aa*) = 1 in 50 = $\frac{1}{50}$ = 0.020

Since the total of all frequencies must equal 1, the frequency of the remaining homozygous dominant genotype, *AA* (which does not express disease nor have carrier status) can be calculated by difference:

Incidence of *AA* = 1 - (0.00040 + 0.020)

= 1 - 0.0204

= 0.9796

To summarize the *observed* frequencies of the three genotypes are:

Genotype	AA	*Aa*	*aa*
Observed frequency	0.9796	0.020	0.00040

Next calculate the expected frequencies if the Hardy-Weinberg equilibrium is operating starting with the frequency of *aa* which is the frequency of the disorder i.e. 1 in 2,500.

Frequency of affected individuals (*aa*) = 0.00040 = q^2

Therefore $q = \sqrt{q^2} = \sqrt{0.00040} = 0.020$

Since $p + q = 1$

$p = 1 - q = 1 - 0.020 = 0.98$

Using these values for p and q the frequencies of the other two genotypes can be calculated:

Frequency of Aa = $2pq$ = 2 x 0.98 x 0.020 = 0.0392

Frequency of AA = p^2 = 0.98^2 = 0.9604

Tabulate this data then calculate X^2:

$$X^2 = \sum (O - E)^2/E$$

Genotype	Frequency		$(O - E)$	$(O - E)^2$	$(O - E)^2/E$
	Observed	***E***xpected			
AA	0.9796	0.9604	0.0192	0.00036864	0.000383840
Aa	0.02	0.0392	-0.0192	0.00036864	0.009404082
aa	0.0004	0.0004	0	0	0
Total:	1.0000	1.0000	0	0.0007372	0.0097878

X^2 is the sum of all the values in the final column = **0.010** (2 sig figs)

Normally the degrees of freedom would be 3 - 1 = 2. However, since one of the observations (frequency of disease) was used to estimate the expected values, a further degree of freedom is lost leaving only one.

From tables, the value for P when X^2 = 0.010 is somewhere between 0.95 and 0.99. Therefore there is no significant difference between the observed and expected frequencies so that **the data fit the Hardy-Weinberg equilibrium**.

12. *The following data were obtained for a digoxin radioimmunoassay employing PEG precipitation of the primary antibody. The assay was performed in duplicate. Calculate the digoxin concentration in the serum sample.*

Sample	***Duplicate cpm***	
	1	***2***
TC	*15,100*	*15,900*
NSB	*320*	*380*
TB	*11,350*	*11,650*
0.2 nmol/L standard	*10,320*	*10,980*
0.4 " "	*9,250*	*8,340*
0.8 " "	*6,782*	*6,630*
1.2 " "	*5,104*	*5,890*
2.4 " "	*3,700*	*3,430*
4.8 " "	*1,350*	*1,650*
Patient serum	*4,350*	*5,000*

Mean *NSB* = (320 + 380)/2 = 350 cpm

Mean *TB* = (11,350 + 11,650)/2 = 11,500 cpm

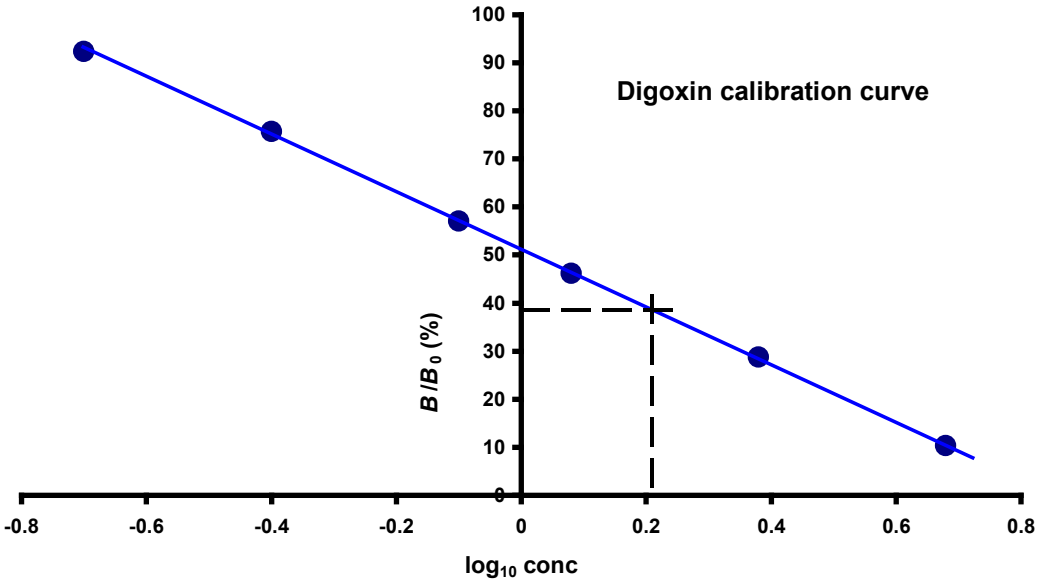
B_0 = Mean *TB* - Mean *NSB* = 11,500 - 350 = 11,150 cpm

Calculate the mean for each pair of duplicates then B/B_0 (%) using the formula:

$$B/B_0\ (\%) = \frac{(\text{Mean cpm standard/sample} - \text{Mean } NSB) \times 100}{B_0}$$

These calculations are performed in the following table:

Sample	Conc	log conc	Duplicate cpm		Mean cpm	Mean – *NSB*	B/B_0 (%)
NSB	-	-	320	380	350		
TB	0	-	11,350	11,650	11,500	1,150	100
Standard	0.2	-0.70	10,320	10,980	10,650	10,300	92.3
"	0.4	- 0.40	9,250	8,340	8,795	8,445	75.7
"	0.8	-0.10	6,782	6,630	6,706	6,356	57.0
"	1.2	0.08	5,104	5,890	5,497	5,147	46.2
"	2.4	0.38	3,700	3,430	3,565	3,215	28.8
"	4.8	0.68	1,350	1,650	1,500	1,150	10.3
Serum	?	?	4,350	5,000	4,675	4,325	38.8



From calibration curve $\log_{10}$ conc when serum B/B_0 (%) = 0.21

Serum digoxin (nmol/L) = antilog$_{10}$ 0.21 = **1.6 nmol/L**

S

T

U

V

W

Z

Allan Deacon BSc PhD FRCPath was, until his retirement in 2006, Consultant Clinical Scientist in the Clinical Biochemistry Department at Bedford Hospital.

Prior to that he held the position of Consultant Clinical Scientist at Kings College Hospital (London) and previously trained in Clinical Biochemistry at Northwick Park Hospital.

His main interest for many years was porphyria and he provided a Supraregional Analytical and Advisory Service for porphyrins.

Published by ACB Venture Publications